universität wien

DISSERTATION

Reasoning about weight in great apes (*Pan troglodytes, Pan paniscus, Gorilla gorilla gorilla, Pongo pygmaeus abelii*) and 3- to -4-year-old children

angestrebter akademischer Grad

Doktor/in der Naturwissenschaften (Dr.rer.nat.)

Verfasserin / Verfasser: Mag. Cornelia Schrauf

Dissertationsgebiet (lt. Studienblatt): Zoologie

Betreuerin / Betreuer: Ao. Univ.-Prof. Dr. Ludwig Huber

Wien, im April 2010

Acknowledgments

This PhD project would not have been realizable without the help and support of many people. At this point, it is a pleasure for me to name some of them.

A very special thank you goes out to my supervisor on site Josep Call, who gave me the opportunity to work with great apes at the Wolfgang Köhler Primate Research Center in Leipzig. He was the first contact person for all kinds of professional questions and helped me to keep on going even when it seemed that nothing would work out. I furthermore want to thank him for all numerous stimulating discussions, as well as his patience and support on all levels. I am very grateful for everything that I learned from him in both scientific and human respects.

I would also like to thank my thesis supervisor Ludwig Huber. Firstly, for his support and for encouraging me to conduct this experiment in Leipzig. Secondly, and very importantly, for always being ready to help and provide me with valuable suggestions. Without his help from the early beginning I would not have reached my current position.

A special thank you goes to Eva Millesi and the members from the Department of Behavioral biology, who took care of me during the phase of writing up my thesis and restored my self-confidence, at times I needed it most.

Importantly, I also want to thank Prof. Bruce Hood and Prof. Antonio Souto for their willingness to review this thesis.

I am forever indebted to my parents for their understanding, endless patience and encouragement when it was most required.

Finally, I would like to express special thanks to my life partner Simo El Hafiane, who helped me to concentrate on completing this dissertation and supported me mentally during the course of this work. Without his help and encouragement, this study would not have been completed.

This research was supported by a grant from the Austrian Academy of Science (DOC fFORTE).

Contents

Synopsis

Understanding relationships between important events separated in time and space can be achieved by two possible ways: associative learning and causality. Associative learning is the cognitively less demanding approach because it depends on an organism experiencing many repetitions of two events in close spatial and temporal contiguity. When cause and effect are separated in space or time, associative learning loses its functionality.

Understanding causality is a cognitively higher-level mechanism as: it requires the individual to understand how the antecedent event A (the cause) produces the consequent event B (the effect) through some mediating force that explains why the first invariably leads to the second. Causal understanding is *a priori* and leads to a flexible approach to problem solving in the physical realm, because individuals may substitute antecedent events in initiating the same mediating force or prevent consequent events by blocking the antecedent´s initiation. Thus, causality allows individuals to transfer what they have learned in one task to a novel one based on the same principles, whereas associative learning results in the formation of a task-specific solution.

Knowledge about physical causality – defined as the way that objects behave and interact – is a central issue in cognitive development (Haith & Benson, 1988; Piaget, 1952; 1954; Spelke & Newport, 1998). Most of the early research on young children´s understanding of causality was highly descriptive, and children´s verbal explanations of real or hypothetical events were used to highlight the differences in causal understanding between children and adults (e.g. Piaget, 1930). More recent studies, using the 'violation of expectancy' technique, reveal that even infants as young as six months are sensitive to causal relations (Leslie, 1982; Leslie & Keeble, 1987; Oakes & Cohen, 1990). The used experimental techniques rely on the amount of time infants spend looking at a certain event. After infants are familiarized with an event sequence, they are presented with two test trials, one that obeys the laws of physics (possible) and others that violate these laws (impossible). Infants' looking at the two test events is then measured. The relevant issue is whether infants will increase their looking times when there is a change of the event from possible to impossible or vice versa. If there is a difference, and it is the impossible event that infants look at longer, it is taken as evidence that they have detected the impossibility of the event sequence. Research using these methods has demonstrated that infants were sensitive to properties of the physical world that were previously assumed to be far beyond their understanding (e.g. Baillargéon & Hanko-Summers, 1990; Spelke, 1988; 1990; 1991). These methods, however, are hotly contested (see Karmiloff-Smith, 1992) and remain an important issue.

The degree to which animals understand physical principles is a matter of contention. Researchers have tried, for many decades, to get a glimpse into the folk physics of animals, i.e. their common sense understanding of how the world works, as well as why it works in the way it does (e.g. Köhler, 1925; Povinelli, 2000; Tomasello & Call, 1997). Concerning primates in particular, several studies have suggested the use of various cognitive strategies in understanding such things as space, tools, object categories and quantities. However, relatively little is known about nonhuman primates´ understanding of the underlying causal mechanisms. To date, what we know about primates' causal understanding is almost exclusively restricted to the context of tool use (see Bard et al. 1995; Limongelli et al. 1995; Hauser, 1997; Natale, 1989; Spinozzi & Potí, 1989; 1993; Visalberghi & Limongelli, 1994; Visalberghi & Trinca, 1989; Visalberghi et al. 1995; Povinelli, 2000). Surprisingly little effort has been made to experimentally determine what exactly these animals understand about the causal mechanisms involved in these interactions.

Causality is the cognitive basis for acquiring and using physical concepts. Some researchers (Penn & Povinelli, 2007; Vonk & Povinelli, 2006; Povinelli, 2004) argued that only humans possess the capability to reason about physical concepts such as gravity, force or weight that refer to unobservable entities. Povinelli (2004) stated that reasoning about weight is beyond the scope of chimpanzees' abilities, as it requires the human language to connect perception of weight with a concept of weight. In his opinion the weight concept is a theoretical construct that cannot be detected visually (i.e. because it is unobservable) and, in contrast to human adults and children, chimpanzees rely strictly upon observable features.

It is true that, in contrast to the wealth of knowledge available on the use of visual information in primates, we know little about their use of other sensory modalities. This is especially true of their kinesthetic abilities, in particular their ability to discriminate weight. To date, only two studies have been published on this topic. McCulloch (1941) trained five chimpanzees to select the heavier of two boxes. Despite the large weight differences (80 vs 480 g or 80 vs. 640 g), the chimpanzees needed a median of 1100 trials to master the task.

Klüver (1933) also conducted weight discrimination experiments with monkeys. Subjects were presented with two boxes differing in weight (450 vs 150 g) and were required to compare the weights and select the correct alternative: some subjects needed to select the heavier box, others the lighter box. Subjects needed hundreds of trials to reach criterion, even though the weight differences between quantities were considerable. The large number of trials needed seems to indicate that subjects had difficulty encoding weight cues. Nonetheless, note that McCulloch´s (1941) and Klüver´s (1933) study are both based on solving an arbitrary relation and not a causal one, the correct solution being decided by the experimenter. Thus, there was actually no *a priori* reason for the tested subjects why the heavy or light object should be correct. It is possible that subjects would more easily learn to use weight information when weight is causally related to solving the task.

Children's understanding of weight is initially grounded in their perceptual sensation of felt weight rather than in a property of matter (Piaget & Inhelder, 1974). Although even preschool children understand that objects vary in weight, their initial concept of weight is the felt pressure of objects on their hands and the effort they have to exert to hold things or to make them move. For instance, more than half of 4-year-olds, and half of 6-year-olds, judged that a small rice corn or a piece of styrofoam weighs nothing at all (Smith, Carey, Wiser, 1985). Children at this age also judge that two visibly different-sized clay balls weigh the same because they both feel light, rather than thinking their weight must differ because they posses different amounts of matter. Thus, younger children's concept of weight is highly influenced by the felt weight of an object, and their reliance on "felt weight" encourages size and density intrusions on weight judgments. Size intrusions on weight judgments are evident in the young child´s thinking that relative size is a reliable indicator of weight. Children at this stage assume that popcorn becomes heavier when it is popped because it got bigger (Piaget & Inhelder, 1974). They will also predict that a wax and clay ball will weigh the same if they were made the same size.

According to Smith et al. (1985), it is around 9-years of age that children develop a concept of weight that is less dependent on their perceptual sensations. While young children at ages 3 to 4 have early theories that contain concepts of weight and size (but totally lack the concept of density), at ages 5 to 7 their concept of weight becomes modified. At this point they can be considered to have an undifferentiated weight/density concept. The change in the child´s concept of weight is indicated by the ability to regard weight differences as important in distinguishing whether objects are made of the same material. Finally, 8- to 9-year-old children reach the point where they not only have distinct density and weight concepts, but also reconceptualize matter as the underlying constituent of objects. Such children assert that everything must weigh something, no matter how small it is, and judge that a small piece of clay is still heavier than a piece of play dough the same size.

Smith et al. (1985) also claim that 3- to- 4-year-olds consider weight as a physical property of objects which causally affects that object's interactions with other objects. This is indicated by their belief that a heavy object will make a foam rubber bridge collapse. The transition from felt weight to weight as a property of matter has been studied in the frame of the balance-scale paradigm (Metz 1993). The literature typically assumes that children´s initial representation of the elements on the balance-scale is framed in terms of weight (e.g. Case 1985; Siegler, 1978). In the experimental procedure on which these conclusions are based, however, an unequal number of identical elements is placed on each end of the scale to achieve differences in weight, rather than presenting visually identical heavy and light objects. Thus, the number prediction pattern is confounded with the weight prediction pattern. It remains open whether younger children understand the relevance of weight to the balance-scale problem, i.e. that heaviness causes the scale to tip.

In the course of my PhD, I investigated great apes´ and 3- to -4-year-old children's understanding of weight as a non-visual-based object property and whether they are genuinely able to reason about weight. To do so, I tested chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), gorillas (*Gorilla gorilla gorilla*) and orangutans (*Pongo pygmaeus abelii*) housed at the Wolfgang Köhler Primate Research Centre, Leipzig Zoo (Germany). The children were tested at the University of Heidelberg, Institute of Psychology.

The first experiment was designed to evaluate how bonobos, gorillas and orangutans perform in a weight discrimination task. To test for a possible superiority of one sensory modality, a second condition, based on visual discrimination (achromatic color), was included as a control. In order to avoid training the subjects to pay attention to the kinesthetic cue, I used the exchange paradigm as the method of choice. In the exchange task, subjects have to give objects to the experimenter to get rewarded; lifting the objects (perceiving the kinesthetic cue) was therefore an integral part of the procedure. This methodology also allowed a direct comparison between visual and weight discrimination.

Twelve subjects (5 bonobos, 2 gorillas and 5 orangutans) were presented with two sets of objects corresponding to two conditions. The objects in the achromatic color condition (white/black) differed only in color, and those in the weight condition (light/heavy) differed only in weight. Five apes learned to discriminate weight and six to discriminate color. Subjects learned color discrimination faster than weight discrimination. The results suggest that bonobos and orangutans are sensitive to differences in weight and are able to learn to discriminate objects that differ in this property.

Based on these findings, I conducted a second experiment to assess whether subjects would be able to use weight information in a situation in which weight is causally related to solving the task. In particular, I presented apes, which had already participated in the weight exchange experiment, with two opaque cups and hid slices of two bananas in one of them. Subjects had the possibility to lift both cups before making a choice. The aim of the study was to test whether great apes would independently gain kinesthetic information about cup weight and then consequentially select the heavier one, i.e. the one containing the banana slices. To rule out that good performance was influenced by previous experience in weight experiments or by the large amount of food they received as reward, I presented the same task to a group of chimpanzees that were naïve to weight experiments. These subjects also participated in an additional test condition in which the same problem was presented but based on learning to associate arbitrary visual stimuli. Specifically, cups of the same weight but different visual appearance were presented. Subjects had to learn to associate a specific visual pattern as an indicator for the location of the reward. Hence, I presented chimpanzees with two task stimuli: one held a causal relation (i.e. perceived weight indicates the hidden food), the other an arbitrary relation (i.e. certain visual pattern indicates food), with the contingencies of reinforcement being the same. The results show that experience did not affect performance because the nine naïve subjects were equally able to find the food in the causal condition, i.e. food inside the cup generated heaviness. Interestingly, only one of the naïve subjects solved the task in the arbitrary condition. Our results confirm previous findings (Call 2004; 2006a, b; 2007; Hanus and Call 2008) that apes perform better in problems grounded on causal compared to arbitrary relations.

Finally, I investigated whether 3- to -4-year-old children are able to correctly encode the relevance of weight in influencing the behavior of a balance-scale. I studied how well 3- to 4-year-old children learn to use one of two different weights (of equal appearance) to tip the scale. In one condition (possible), the heavy weight produced the desired outcome; in the other condition (impossible), the light weight caused the scale to tip. Only 4-year-olds learned more effectively in the possible versus impossible condition, suggesting that children younger than 4 have not yet developed clear expectations of the role of weights on the movements of a balance-scale.

In summary, I found that apes are able to discriminate weight and can quickly learn to use weight information in a situation in which weight is causally related to solving a task. In contrast to Povinelli´s (2004) claims, the human language does not seem to be essential to reason about weight. Although I have not tested explicitly for the prerequisites for successfully completing the presented tasks my experiments suggest that it is not the human language that allows to reason about unobservable entities. Chimpanzees and other apes do not rely strictly on observable features. In contrast, performance seemed to depend on whether the task was presented in a causal compared to an arbitrary context. In children, performance seemed to be age dependent because only the 4-year-olds showed a difference in learning between the condition that obeyed the laws of physics and the condition that violated these laws.

References

Baillargéon R, Hanko-Summers S (1990) Is the object adequately supported by the bottom object? Young infants´ understanding of support relations. Child Development 5, 29-54

Bard KA, Fragaszy D, Visalberghi E (1995) Acquisition and comprehension of tool-using behavior by young chimpanzees (*Pan troglodytes*): effects of age and modeling. International Journal of Comparative Psychology 8, 1-22

Call J (2004) Inferences about the location of food in the Great Apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla, and Pongo Pygmaeus*). Journal of Comparative Psychology 115, 159-171

Call J (2006a) Descartes´ two errors: Reason and reflection in the great apes. In: Hurley S, Nudds M (eds) Rational animals (pp.219-234). Oxford: Oxford University Press

Call J (2006b) Inferences by exclusion in the great apes: the effect of age and species. Animal Cognition 9:393-403

Call J (2007) Apes know that hidden objects can afford the orientation of other objects. Cognition 105: 1-25

Case R (1985) Intellectual development: Birth to adulthood. New York: Academic Press

Haith MM, Benson J (1988) Infant cognition. In: Kuhn D, Siegler RS (eds) Handbook of Child Psychology. Vol 1: Cognition, perception and language (pp. 199-254). New York: Wiley

Hanus D, Call J (2008) Chimpanzees infer the location of a reward based on the effect of its weight. Current Biology 18, R370-R372

Hauser MD (1997) Artifactual kinds and functional design features: What a primate understands without language. Cognition 64, 285-308

Karmiloff-Smith A (1992) Beyond modularity. A developmental perspective on cognitive science. Cambridge, MA: MIT Press

Klüver H (1933) Behavior mechanisms in monkeys. Chicago University Press, Chicago

Köhler W (1925) The Mentality of Apes. New York: Liveright

Kremers J (2005) The primate visual system: A comparative approach. John Wiley and Sons

Leslie AM (1982) The perception of causality in infants. Perception 11, 173-186

Leslie AM, Keeble S (1987) Do six-month-old infants perceive causality? Cognition 25, 265-288

Limongelli L, Boysen ST, Visalberghi E (1995) Comprehension of cause-effect relations in a tool-using task by chimpanzees (*Pan troglodytes*). Journal of Comparative Psychology 109, 18-26

McCulloch T (1941) Discrimination of lifted weights by chimpanzees. Journal of Comparative Psychology 32, 507-519

Metz KE (1993) Preschoolers´ Developing Knowledge of the Pan Balance: From New Representation to Transformed Problem Solving. Cognition and Instruction 11, 31-93

Natale F (1989) Causality II. The stick problem. In: Antinucci F (Ed) Cognitive structure and development in nonhuman primates (pp 121-133). Erlbaum, Hillsdale, NJ

Oakes LM, Cohen LB (1990) Infant perception of a causal event. Cognitive Development 5, 193-207

Penn D, Povinelli D (2007) Causal cognition in human and nonhuman animals: A comparative, critical review. Annual Review of Psychology 58, 97-118

Piaget J (1930) The child´s concept of physical causality. London, K. Paul, Trench, Trubner & Co

Piaget J (1952) The origins of intelligence in children. New York: International University Press

Piaget J (1954) The construction of reality in the child. Ballantine Books, New York

Piaget J, Inhelder B (1974) The child's construction of quantities. London: Routledge and Kegan Paul

Povinelli D (2000) Folk Physics for Apes, Oxford, Oxford University Press

Povinelli D (2004) Behind the ape´s appearance: escaping anthropocentrism in the study of other minds. Daedalus Winter Vol, 133, No.1: 29-41

Siegler RS (1978) Children´s thinking: What develops? Hillsdale, NJ: Lawrence Erlbaum Associates Inc

Smith C, Carey S, Wiser M (1985) On differentiation: a case study of the development of the concepts of size, weight and density. Cognition 21, 177-237

Spelke ES (1988) The origins of physical knowledge. In: Weiskrantz L (Ed) Thought without language. Oxford, UK: Oxford Press

Spelke ES (1990) Principles of object perception. Cognitive Science 14, 29-56

Spelke ES (1991) Physical knowledge in infancy: Reflections on Piaget´s theory. In: Carey S, Gelman R (eds) The epigenesis of mind: Essays on biology and cognition (pp 133-169). Hillsdale, NJ: Lawrence Erlbaum Associates Inc

Spelke ES, Newport EL (1998) Nativism, empiricism and the development of knowledge. In: Learner RE (Ed) Handbook of child psychology: Theoretical models of human development. Vol 1, New York: Wiley

Spinozzi G, Potí P (1989) Causality I: The support problem. In: Antinucci F (Ed) Cognitive structure and development in nonhuman primates (pp 113-119). Erlbaum, Hillsdale, NJ

Spinozzi G, Potí P (1993) Piagetian stage 5 in two infant chimpanzees (*Pan troglodytes*): the development of permanence of objects and the spatialization of causality. International Journal of Primatology 14, 905-917

Tomasello M, Call J (1997) Primate cognition. Oxford University Press, Oxford

Visalberghi E, Trinca L (1989) Tool use in capuchin monkeys: distinguishing between performing and understanding. Primates 30, 511-521

Visalberghi E, Limongelli L (1994) Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). Journal of Comparative Psychology 108, 15-22

Visalberghi E, Fragaszy DM, Savage-Rumbaugh S (1995) Performance in a tool-using task by common chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), an orangutan (*Pongo pygmaeus*), and capuchin monkeys (*Cebus apella*). Journal of Comparative Psychology 109, 52-60

Vonk J, Povinelli DJ (2006) Similarity and difference in the conceptual systems of primates: The unobservability hypothesis. In: Wasserman E, Zentall T (eds) Comparative Cognition: Experimental Explorations of Animal Intelligence (pp 363-387). Oxford University Press.

Great apes` performance in discriminating weight and achromatic color

Cornelia Schrauf[1], Josep Call[2]

1 University of Vienna, Department of Cognitive Biology

2 Max Planck Institute for Evolutionary Anthropology, Leipzig

Published in:

Animal Cognition (2009) 12, 567-574

Abstract

Much work has been done on visual discrimination in primates over the past decade. In contrast, very little is known about the relevance of non-visual information in discrimination learning. We investigated weight and achromatic color (color, henceforth) discrimination in bonobos, gorillas and orangutans, using the exchange paradigm in which subjects have to give objects to the experimenter in order to receive a reward. Unlike previous studies, subjects were not trained to lift objects because lifting the objects was an integral part of the exchange procedure. This methodology also allowed us a direct comparison between visual and weight discrimination. We presented 12 subjects (5 bonobos, 2 gorillas and 5 orangutans) with two sets of objects corresponding to two conditions. The objects in the color condition (white/black) differed only in color and those in the weight condition (light/heavy) differed only in weight. Five apes learned to discriminate weight and six to discriminate color. Subjects learned color discrimination faster than weight discrimination. Our results suggest that bonobos and orangutans are sensitive to differences in weight and able to learn discriminating objects that differ in this property.

Introduction

Foraging efficiently depends to a great extent on the ability to acquire and process information about environmental stimuli. The senses are a main way to acquire this information to make efficient foraging decisions (Gibson 1966, 1979; Gibson KR 1986). Moreover, repeated exposure to the same stimuli and their consequences foster learning, which in turn, increase the probability of making optimal foraging decisions.

Primates have traditionally been considered to be primarily visually oriented animals (Fobes and King 1982), relying on their highly developed visual system to optimize their foraging efficiency and avoid toxic substances. However, the olfactory, auditory and proprioceptory senses also play an important role in acquiring relevant information for food selection (Dominy et al. 2001). This seems particularly important considering that a variety of plants and fruits change color or other visual features only a little if at all, during ripening or maturation (Nagy et al. 1980; Gautier-Hion et al. 1985). In fact, in some primate species vision plays only a secondary role in food acquisition. For example, the aye-aye finger taps wood with its finger to produce auditory information in order to locate cavities under the surface (Erickson 1991). Wild-tufted capuchin monkeys locate frogs hiding inside the *guadue* plant by listening for particular sounds (Izawa 1978). Nut-cracking chimpanzees and capuchin monkeys have to rely on non-visual cues (smell, weight, sound) in order to assess the state of a nut, because the protective shell prevents visual assessment of its contents (for a more detailed description, see Visalberghi and Néel 2003). As cracking open a nut is a costly behavior, both in terms of time and energy, it seems worthwhile to open a nut only when it contains the energy-rich kernel inside and not when the kernel has dried (Boesch and Boesch 1983; Boinski et al. 2003). Therefore, it seems reasonable to assume that primates would use a variety of sensory modalities to evaluate and learn about potential food sources.

We know a fair amount about vision as a way to identify stimuli associated with food. Visual acuity and color perception are well developed in primates (Spence 1934; Grether 1940 a, b; Tigges 1963; Tigges and Tigges 1965; Davis and Markowitz 1978; Barbiers 1985; Vonk 2003; Buchanan-Smith 2005).

Furthermore, numerous studies have investigated visual discrimination learning in primates using cues such as color, brightness, shape and size (Jarvik 1953; Fobes and King 1982; Rumbaugh and Pate 1984). Naïve individuals typically require dozens of trials to associate a particular object with the presence of a reward, although other studies show that individuals can detect the presence of novel objects in their enclosure quite readily (Menzel and Juno 1982). Additionally, most studies investigating problem solving of various types (e.g., tool-use, inferences, computerized tasks) have relied on the use of visual information (Tomasello and Call 1997; Matsuzawa et al. 2006). Even those studies that have investigated whether individuals can make inferences about food location based on the notion of weight have relied on visual information to present the task (Hanus and Call 2008).

In contrast to the wealth of knowledge available on the use of visual information in primates, we know much less about the use of other sensory modalities. This is especially true of their kinesthetic abilities, for instance, with regard to weight discrimination. To the best of our knowledge, there are only two published studies in this area. McCulloch (1941) trained five chimpanzees to select the heavier of two boxes. Despite the large weight differences (80 vs. 480g or 80 vs. 640g), the chimpanzees needed a median of 1,100 trials to master the task. Once the subjects were successful in their performance, the weight of the heavier box was decreased. The smallest weight difference that the chimpanzees were able to discriminate between was 4 g (80 vs. 84g).

Klüver (1933) also conducted weight discrimination experiments with monkeys. Subjects were presented with two boxes differing in weight (450 vs. 150g) and were required to compare their weights and select the correct alternative: some subjects needed to select the heavier box while others were required to select the lighter box. Long-tailed macaques and spider monkeys took a median of 586 and 330 trials, respectively, to reach an errorless performance whereas a capuchin monkey required 321 trials. Moreover, species varied in the weight difference between quantities that they required in order to see signs of successful discrimination (long-tailed macaques: 750g; spider monkeys: 900g; capuchin monkey: 600g). Thus, these two studies indicate that subjects of various species are able to discriminate weight, but it took them hundreds of trials to reach criterion even though there were considerable weight differences between quantities.

One might conclude, based on this evidence, that visual discrimination is easier than weight discrimination, thus reinforcing the idea of a visual-sensory primacy in monkeys and apes. However, caution is warranted when comparing experiments on visual and weight discrimination because they used different methods. Additionally, only a total of 11 individuals (belonging to 4 different species) have been tested on weight discrimination compared to hundreds of individuals from dozens of species that have been tested on visual discrimination. It is, therefore, essential to study the use of non-visual information to get a more balanced view of discrimination learning with non-visual stimuli. Such information can be used to design more complex problem-solving tasks that do not rely solely on visual input (e.g., Povinelli and Dunphy-Lelii 2001) and to gain some insights about the use of weight to assess the suitability of stone hammers to crack open nuts (Boesch and Boesch 1983; Visalberghi et al. 2007; Schrauf et al. 2008).

The aim of the current study was to evaluate how great apes, other than chimpanzees, perform in a weight discrimination task. To date nothing is known about the ability of gorillas, orangutans and bonobos to use weight information. For this purpose, we employed a method in which, unlike previous studies (Klüver 1933; McCulloch 1941; Schrauf et al. 2008), subjects lifted the objects spontaneously, without the need to offer an incentive to do so. Furthermore, the methodology used here allowed us to directly compare visual discrimination in the form of an achromatic color (color, henceforth) discrimination task. This means that we were able to assess the putative superiority of visual over kinesthetic information. Additionally, the inclusion of this visual discrimination task served as a control for the procedure because we expected that subjects would be able to solve at least the color discrimination task within the allocated number of trials.

Method

Subjects

Five orangutans (*Pongo pygmaeus abelii*), three gorillas (*Gorilla gorilla gorilla*) and five bonobos (*Pan paniscus*) housed at the Wolfgang Köhler Primate Research Centre, Leipzig Zoo (Germany) participated in this study. There were 8 females and 4 males with ages ranging from 11 to 33 years (for a more detailed description, please refer to Table 1). All subjects lived in social groups of various sizes in indoor and outdoor compounds. Prior to this study, subjects had participated in a number of cognitive experiments including visual discrimination of shape, size, and color (e.g., Vlamings et al. 2006) but had never been tested on weight discrimination. The selection of subjects was based on their preexisting experience in exchanging objects for rewards with a human experimenter (see Pele et al. 2009), which represented an integral part of our current procedure. During the experiment, the subjects were separated from their group members and individually tested in their indoor cages.

Materials

Subjects were presented with two sets of objects. For the weight discrimination condition, we used 12 grey PVC cylinders (9.5 cm x 4 cm) of identical outward appearance. Six of the cylinders were filled with lead shot, weighing a total of 400g. The other remaining six were empty, and each weighed 100g. A transparent silicone paste was added to incorporate the contents in a solid and homogenous mass, preventing any rattling noise during manipulations and evenly distributing the weight throughout the tool. For the color condition, we presented 12 colored PVC rectangles (8 cm x 3.5 cm x 1.5 cm). Six of the rectangles were black and the remaining six were white. The objects were placed inside the subjects´ cage in alternate order (see Fig. 1). A blue and a white bowl were placed in front of the cage on the ground. We used grapes as rewards for all the subjects except Bimbo, a male orangutan, who received monkey chow instead.

Figure 1

Procedure

The experiment was conducted within the framework of the exchange paradigm (Hyatt and Hopkins 1998) in which subjects had to exchange objects with a human experimenter. In our study, the apes had to return objects with a certain feature to the experimenter to obtain a food reward. The advantage of the exchange paradigm was that the act which was needed in order to assess the critical property (lifting of the object) was a part of the procedure. In particular, giving the experimenter the objects that were placed inside the subjects' cage necessarily involved lifting them prior to the exchange. Thus neither an incentive nor training was needed for the subjects to lift the objects in order to perceive kinesthetic cues.

Before the apes entered the testing room the objects were positioned inside the subjects' cage on a ledge by one of the research panels (see Fig. 1). The experimenter was sitting outside the cage facing the subject. To exclude inadvertent cueing while the subject was selecting the objects, an opaque panel prevented the experimenter from seeing the subjects. Only after the subject had handed the selected item to the experimenter, and thus had already made its choice, the experimenter could see or feel if it was the correct one. This means that apes could not receive experimenter-given cues regarding the objects' value. Below the opaque panel, a wire mesh divider allowed the exchange of objects for rewards between the experimenter and the subject. The experimenter requested an item by extending her hand, palm-up (begging gesture), in front of the wire mesh. If the subject returned a correct item, the experimenter dropped it in the white bowl on her right and gave a reward to the subject. If the subject returned an incorrect object the experimenter dropped it in the blue bowl on her left and gave no reward.

Subjects were presented with two sets of objects corresponding to two conditions. In the color condition (white/black) the items differed only in achromatic color; in the weight condition (heavy/light) they differed only in weight. Prior to testing, we decided which feature (black or white; heavy or light) would be regarded to be correct for each subject. This selection was done arbitrarily e.g., there was no obvious reason for the apes why black or white (heavy or light) should be correct and counterbalanced across subjects.

Each subject received two sessions per day. A session lasted until all six correct objects had been returned to the experimenter. Everyday after the first session, the ape was moved to an adjacent cage to allow the experimenter to reposition the objects for the following session. Subjects were tested until they either reached the criterion for errorless performance (see below) or had received 36 sessions.

Data scoring and analysis

We videotaped all the trials and scored them live on a check-sheet. We used the videotapes to verify the original scoring. We coded the identity of the objects returned by the subjects. The criterion for successful performance consisted of returning all six correct objects in their first six attempts during a given session (errorless performance). As the probability for such a result to occur when the objects are returned randomly is 0.001 for a single session and the combined probability for the whole test of 36 sessions is 0.039, we considered this criterion as sufficient.

We used non-parametric statistics to analyze the data. In particular, we compared the number of subjects that solved each task (sign test) and the number of sessions needed to discriminate color and weight (Wilcoxon test). Additionally, we assessed the effect of the order of task presentation and the type of stimuli on performance (Fisher's test). Finally, we tested whether bonobos and orangutans differed in weight and color discrimination (Mann-Whitney-U test). Gorillas were not included in the species comparison because of our small (N=2) sample size. All tests were two-tailed.

Results

Five of the 12 subjects tested performed above chance in the weight discrimination condition (see Table 1). They needed a median of 30 sessions (331 exchanges) to reach the criterion of 6 out of 6 correct exchanges. Six of the 12 apes solved the color discrimination with a median of 6.5 sessions (64 exchanges). The remaining subjects showed no improvement in performance until the end of the experiment (session 36). Only three subjects learned to discriminate both color and weight. A comparison of the number of subjects that solved color versus weight revealed no significant difference using the sign test ($N = 3$; $P > 0.99$). However, when all subjects were taken together as a group they seemed faster at discriminating color than weight (Wilcoxon test $Z = -2.033$; $N = 12$; $P < 0.042$).

Table 1

Subjects learned the color discrimination in both cases - when white or black was the positive stimulus. Likewise weight discrimination was also learned on both instances independently if heavy or light was regarded to be correct. Neither the type of positive stimulus nor the order of presentation of conditions had a significant influence on the subjects´ performance in either the color discrimination task (Fisher exact tests: type, $P > 0.99$; order, $P > 0.99$) or the weight discrimination task (Fisher exact tests: type, $P = 0.072$; order, $P > 0.99$). There were no significant differences in performance between bonobos and orangutans in either the color (Mann-Whitney U test $U = 4$; $N = 10$; $P = 0.058$) or weight discrimination conditions (Mann-Whitney U test $U = 9$; $N = 10$; $P = 0.435$). Because of the small sample size (N = 2), the gorillas were not included in the analysis for species differences. One gorilla (Bebe) learned to discriminate color within seven sessions. However, in the weight discrimination condition she showed no improvement in performance until the end of the experiment (36 sessions).

Discussion

Some of our bonobos and orangutans learned to discriminate between differences in weight; however, they learned color faster than weight discrimination. We found no evidence of significant differences between bonobos and orangutans concerning weight discrimination. None of the two gorillas learned the weight discrimination task but the small sample size precluded a direct comparison with the other species.

It may seem surprising that most of the subjects did not master the weight or color discrimination task. A lack of motivation for this outcome seems unlikely because even unsuccessful subjects continued to return objects at every opportunity and yet showed neither an increase nor a decrease in their performance over the course of the experiment. Similarly, a failure to understand the task is also unlikely because all subjects had been fully trained in the exchange procedure prior to the current study. Perhaps this result is not so surprising if one considers that learning to associate an arbitrary cue such as color or weight with a reward is not a trivial task for great apes (Call 2006). Naïve apes may take dozens of trials (if not more) to master a seemingly easy task such as choosing between two alternatives presented on a platform (Spence 1937; Jarvik 1953; Rumbaugh and Rice 1962; Davis and Markowitz 1978).

Compared to previous studies, orangutans and bonobos required a much smaller number of sessions to reach the weight discrimination criterion than chimpanzees and monkeys (Klüver 1933; McCulloch 1941). This difference is also true when comparing the number of exchanges / trials in both sets of studies (median of 331 exchanges in our study/ median of 1,100 trials in Mc Culloch´s experiment). A direct comparison with the monkeys' performance tested by Klüver is problematic because the author enlarged the weight differences until the first sign of learning occurred; however, the monkeys needed a median of 414.5 trials to reach an errorless performance. One possible explanation for this difference is that the species differed in their ability to discriminate weight. Alternatively, differences in the method employed in order to assess weight discrimination could explain these differences. In particular, the exchange method used in the current study meant that no training was required to lift the objects –subjects did so within the context of the task; thus, the apes´ task consisted of perceiving and associating the discriminative feature that was associated with reinforcement. One final explanation has to do with inter-individual differences, which were important in our study. Such large inter-individual differences paired with small sample sizes could potentially explain some of the differences observed between studies. It is currently unclear which of these three alternatives is responsible or may have contributed to a greater degree to the observed differences. Future studies based on the exchange method, comparing the primate species included in previous studies, will be needed in order to tease apart these alternatives.

We found some evidence that subjects were faster at discriminating color than weight. Unlike previous studies, this difference cannot be attributed to the use of different methodologies to test each perceptual modality. This would seem to reinforce the notion of a primacy of visual over kinesthetic information, at least in the case of bonobos and orangutans. At this point it is important to note that all of our apes were naïve to weight experiments but had participated earlier in experiments involving visual discrimination. This difference might therefore have influenced the subjects' performance and it remains possible that apes would have been faster when given experience with weight before. Therefore, it is conceivable that the putative superiority of visual over kinesthetic cues reflected an effect of practice.

A remarkable observation from the experiment should be mentioned. All subjects that were able to solve the weight discrimination condition started to divide the objects according to its weight. In particular, whenever an incorrect item was chosen (touched with the finger or lifted), the subjects placed it on the ground in front of them. On one instance, a female orangutan placed all the incorrect items in front of her on a piece of cardboard that she had brought with her before entering the testing room. Because an object's weight is only perceived by lifting it in this study, this behavior could have been a strategy to help keep track of the objects that were not rewarded, thereby allowing them to distinguish those objects more readily from the others. On the occasions when the subjects had also placed correct items on the ground, we observed that they would lift each object off the ground until the correct item was found. However, this behavior became less frequent over time.

Although apes can sort objects into distinct groups if trained to do so (Premack 1976; Garcha and Ettlinger 1979; Savage-Rumbaugh 1986; Matsuzawa 1990; Spinozzi 1993; Tanaka 1995), they rarely do so spontaneously (Tomasello and Call 1997, but see Hayes and Nissen 1971 for a notable exception on an enculturated chimpanzee). However, in the present study the apes - without any specific sorting training or special rearing history - spontaneously separated items spatially according to weight. It is highly possible that they did so to keep track of the objects' identity in the absence of any discriminative visual cues. This explanation is further reinforced by a similar finding on hammer selection by a capuchin monkey (Schrauf et al. 2008). This individual lifted the potential hammers, which were similar in outward appearance, and placed the heavier (most efficient) one close to him, often holding it with his foot. In contrast, the lighter tools were placed either back or sideways away from him. Schrauf et al. (2008) concluded that this advantageous strategy allowed the capuchin monkey to keep track of the heavy tool. Coming back to the current experiment, it would be interesting to see whether apes would spontaneously make different piles of visually undistinguishable objects based on their weight while the experimenter was absent, in anticipation of her returning at a later time. This would not only show that apes sort objects spontaneously when the situation requires them to do so but that they could also plan for the occurrence of future events (e.g., the arrival of the experimenter) - something that has been recently reported in the great apes (Mulcahy and Call 2006; Dufour and Sterck 2008; Osvath and Osvath 2008) and corvids (Raby et al. 2007). This form of planning involving the execution of a number of actions in preparation for a particular goal has been described as planning for action. An analogous situation has been reported in dolphins gathering multiple weights and placing them at a particular location in preparation to solve a task (Kuczaj and Walker 2006). Chimpanzees and capuchin monkeys have also been observed to bring nuts and tools to an anvil to crack open the nuts (Boesch and Boesch 1984; Visalberghi et al. 2007).

Perhaps one of the most important contributions of this study is that it calls attention to the use of kinesthetic information in testing cognition in primates and other animals. This is important because most studies have relied on tasks with a predominant visual component and less attention has been given to non-visual modalities when presenting problems. This is even more remarkable given the reported observations on nut-cracking in chimpanzees and capuchins who use weight as a key feature for the selection of hammers to crack open nuts (Boesch and Boesch 1983; Visalberghi et al. 2007; Schrauf et al. 2008). Bril et al. (2008) reported that captive chimpanzees changed their nut-cracking behavior according to the weight of the hammers provided to them. Similarly, Biro et al. (2006) found that wild chimpanzees' choice to use a stone as an anvil or a hammer was affected by its weight. Our study has shown that utilizing weight information is readily done by bonobos and orangutans, and that is picked up much faster than has been shown in previous studies. This is the prerequisite for other tasks such as hammer selection in which weight is a functional feature. An important distinction has to be made between the relation between weight and reward delivery in discrimination tasks (like the current one) and the hammer selection tasks (Call 2006). In discrimination tasks, weight and reward hold an arbitrary relation i.e., there is no reason (other than the experimenter's choice) why 'heavy' should be associated with reward delivery. In contrast, in hammer selection tasks, there is a physico-causal reason (having to do with force) as to why heavy hammers are more effective than lighter ones. Detecting weight differences between stimuli is the first step and the prerequisite to presenting more complex problem-solving tasks based on a richer causal structure.

The distinction between arbitrary and causal relations between stimuli is not restricted to weight appreciation (Hanus and Call 2008) but it extends to other domains including a naïve appreciation of solidity, gravity, and collision (Call 2004; 2007; Martin-Ordas and Call 2009; Martin-Ordas et al. 2008; Bania et al. 2009). Most of these studies are based on visual information but it is conceivable that future studies will document the arbitrary versus causal relation in the kinesthetic modality. For instance, one could investigate whether subjects could infer the presence of a reward in one of two opaque cups based on its weight. Provided that both cups are identical and the reward is placed in one of the cups, the heavier of the two cups should hold the reward. One would predict that subjects would be much faster at solving tasks with a causal structure than an arbitrary one. Recently, Hanus and Call (2008) showed that chimpanzees were able to infer the location of a reward based on the effect that its weight had on a balance beam. Nevertheless, in this study the information about the effect of weight was visual rather than kinesthetic.

The use of information about weight may not be the only kinesthetic information that subjects may successfully use to solve complex problems. Seed (2007) presented five chimpanzees with functional and non-functional trap tubes. One subject, Alexandra, developed a strategy in which she inspected both traps of the apparatus with her fingers before responding. She started to do so after an incorrect response and continued this behavior after every error. Interestingly, Alexandra was the only subject that could solve the task. It is therefore conceivable that the chimpanzee's good performance was influenced by the acquisition of tactile information. Thus, although monkeys and apes are primarily visual animals, other senses (e.g., tactile) also play a major role to gain relevant information from the environment and can be a suitable channel to pose some questions about complex problem solving.

In conclusion, the current study demonstrated that bonobos and orangutans were sensitive to differences in weight between two otherwise identical objects. We think that this kinesthetic ability is of importance as the animals might apply such knowledge to different situations such as tool use and tool production or when making inferences about the physical environment. This study also showed the importance of the employed method to measure a certain variable. Whereas in previous studies the chimpanzees needed approximately 1,000 trials (McCulloch 1941) to reach criterion, the apes in our study learned much faster, although all of them were naïve to weight experiments. The advantage of the exchange paradigm was that the act needed to assess the critical property, namely, lifting the object, was already part of the exchange routine. Further studies in different contexts are needed for a better understanding as to what extent apes can make use of weight information and how this guides their problem-solving behavior.

Acknowledgments

This research was supported by a grant (DOC fFORTE) from the Austrian Academy of Science to Cornelia Schrauf. We thank Bernhard Voelkl for statistical advice and helpful comments on the manuscript, Nathan Pyne-Carter for improving the English and Sandra Michaelis for her drawing of the experimental set-up. The reported experiments comply with all laws of the country in which they were performed.

References

Bania AE, Harris S, Kinsley HR, Boysen ST (2009) Constructive and deconstructive tool modification by chimpanzees (*Pan troglodytes*). Anim Cog 12: 85-95

Barbiers RB (1985) Orangutans´ color preference for food items. Zoo Biol 4: 287-290

Biro D, Sousa C, Matsuzawa T (2006) Ontogeny and cultural propagation of tool use by wild chimpanzees at Bossou, Guinea: case studies in nut cracking and leaf folding. In: Matsuzawa T, Tomonaga M, Tanaka M (eds) Cognitive development in chimpanzees. Springer, Tokyo, pp 476 – 508

Boesch C, Boesch H (1983) Optimization of nut cracking with natural hammers by wild chimpanzees. Behaviour 83: 265-286

Boesch C, Boesch H (1984) Mental map in wild chimpanzees: an analysis of hammer transports for nut cracking. Primates 25: 160-170

Boinski S, Quatrone RP, Sughrue K, Selvaggi L, Henry M, Stickler CM, Rose LM (2003) Do brown capuchins socially learn foraging skills? In: Fragaszy DM, Perry S (eds) The biology of traditions: Models and evidence. Cambridge University Press, Cambridge, pp 365–390

Bril B, Dietrich G, Foucart J, Fuwa K, Hirata S (2008) Tool use as a way to assess cognition: how do captive chimpanzees handle the weight of the hammer when cracking a nut? Anim Cog DOI 10.1007/s10071-008-0184-x

Buchanan-Smith HM (2005) Recent advances in color vision research. Am J of Primatol 67: 393-398

Call J (2004) Inferences about the location of food in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus*). J Comp Psychol 118: 232-241

Call J (2006) Descartes' two errors: reasoning and reflection from a comparative perspective. In: Hurley S, Nudds M (eds) Rational animals. Oxford University Press, Oxford, pp 219-234

Call J (2007) Apes know that hidden objects can affect the orientation of other objects. Cognition 105: 1-25

Davis RR, Markowitz H (1978) Orangutan performance on a light-dark reversal discrimination in the zoo. Primates 19: 755 – 759

Dominy NJ, Lucas PW, Osorio D, Yamashita N (2001) The sensory ecology of primate food perception. Evol Anthropol 10: 171 - 186

Dufour V, Sterck EHM (2008) Chimpanzees fail to plan in an exchange task but succeed in a tool-using procedure. Behav Processes 79: 19-27

Erickson CL (1991) Percussive foraging in the aye-aye, *Daubentonia madagascariensis*. Anim Behav 41: 793-801

Fobes JL and King JE (eds) (1982) Primate behavior. New York Academic Press, New York

Garcha HS, Ettlinger G (1979) Object sorting by chimpanzees and monkeys. Cortex 15: 213-224.

Gautier-Hion A, Duplantier JM, Quris R, Feer F, Sourd C, Decoux JP, Dubost G, Emmons L, Erard C, Hecketsweiler P, Moungazi A, Roussilhon C, and Thiollay JM (1985) Fruit characters as a basis of fruit choice and seed dispersal in a tropical forest vertebrate community. Oecologia (Berl) 65:324-337

Gibson JJ (1966) The senses considered as perceptual systems. Houghton Mifflin, Boston

Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, Boston

Gibson KR (1986) Cognition, brain size and the extraction of embedded food resources. In: Else JG, Lee PC (eds) Primate ontogeny, cognition, and social behaviour. Cambridge University Press, Cambridge, pp 93–103

Grether WF (1940a) A comparison of human and chimpanzee spectral hue discrimination curves. J Exp Psychol 26: 394-403

Grether WF (1940b) Chimpanzee color vision. I. Hue discrimination at three spectral points. J Comp Psychol 29: 167-192

Hanus D, Call J (2008) Chimpanzees infer the location of a reward based on the effect of its weight. Curr Bio 18: 370 – 371

Hayes KJ, Nissen CH (1971) Higher mental functions of a home-raised chimpanzee. In: Schrier AM, Stollnitz F (eds) Behavior of nonhuman primates. New York: Academic Press, pp 59-11

Hyatt CW, Hopkins WD (1998) Interspecies object exchange: Bartering in apes? Behav Proc 42: 177-187

Izawa K (1978) Frog-eating behaviour of wild-capped capuchin (*Cebus apella*). Primates 19: 633-642

Jarvik ME (1953) Discrimination of colored food and food signs by primates. J Comp Psychol 46: 390-392

Klüver H (1933) Behavior Mechanisms in Monkeys. Chicago University Press, Chicago

Kuczaj SA II, Walker RT (2006) How do dolphins solve problems? In: Wasserman EA, Zentall TR (eds) Comparative Cognition: Experimental Explorations of Animal Intelligence. New York: Oxford University Press, pp 580-600

Martin-Ordas G, Call J (2009) Assessing generalization within and between trap tasks in the great apes. Int J Comp Psychol (in press)

Martin-Ordas G, Call J, Colmenares F (2008) Constructive and deconstructive tool modification by chimpanzees (*Pan troglodytes*). Anim Cog 11: 423-430

Matsuzawa T (1990) Spontaneous sorting in human and chimpanzee. In: Parker ST, Gibson KR (eds) Language and intelligence in monkeys and apes. Cambridge University Press, Cambridge

Matsuzawa T , Tomonaga M, Tanaka M (2006) Color Recognition in Chimpanzees (*Pan troglodytes*). In: Matsuzawa T, Tomonaga M, Tanaka M (eds) Cognitive development in chimpanzees. Springer, Tokyo

McCulloch (1941) Discrimination of lifted weights by chimpanzees. J Comp Psychol 32: 507-519

Menzel Jr EW, Juno C (1982) Marmosets (*Saguinus fuscicollis*): Are learning sets learned? Science 217: 750-752

Mulcahy NJ, Call J (2006) Apes save tools for future use. Science 312: 1038-1040

Nagy S, Shaw PE and Wardowski WF (eds) (1980) Fruits of Tropical and Subtropical origin: Composition, Properties and Uses. Florida Science Source Inc., Lake Alfred, Florida, USA, 302-321

Osvath M, Osvath H (2008) Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: self-control and pre-experience in the face of future tool use. Anim Cog 11: 661-674

Pelé M, Dufour V, Thierry B, Call J (2009). Token transfers among great apes: species differences, gestural requests and reciprocal exchange. J Comp Psychol (in press)

Povinelli DJ, Dunphy-Lelii S (2001) Do chimpanzees seek explanations? Preliminary comparative investigations. Canad J Exp Psychol 55:2, 187-195

Premack D (1976) Intelligence in ape and man. Hillsdale, NJ: Lawrence Earlbaum Associates

Raby CR, Alexis DM, Dickinson A, Clayton NS (2007) Planning for the future by western scrub-jays. Nature 445: 919-921

Rumbaugh DM, Rice CP (1962) Learning-set formation in young great apes. J Comp Phys Psychol 55: 866 – 868

Rumbaugh DM, Pate JL (1984) Primates` learning by levels. In: Greenberg G and Tobach E (Eds) Behavioral evolution and integrative levels. Hillsdale, NJ: Lawrence Erlbaum Associates, pp 221-240

Savage-Rumbaugh ES (1986) Ape language. New York: Columbia University Press

Schrauf C, Huber L, Visalberghi E (2008) Do capuchin monkeys use weight to select hammer tools? Anim Cog 11: 413-422

Seed A (2007) Cognition in rooks and chimpanzees: A case of convergent evolution? Ph.D. thesis, Cambridge University, Cambridge

Spence KW (1934) Visual acuity and its relation to brightness in chimpanzee and man. J Comp Psychol 18: 333-361

Spence KW (1937) Analysis of the formation of visual discrimination habits in chimpanzee. J Comp Psychol 23: 77-100

Spinozzi G (1993) Development of spontaneous classificatory behaviour in chimpanzees (*Pan troglodytes*). J Comp Psychol 107: 193-200

Tanaka M (1995) Object sorting in chimpanzees (*Pan troglodytes*): classification based on physical identity, complementarity and familiarity. J Comp Psychol 109: 151-161.

Tigges J (1963) On color vision in gibbon and orang-utan. Folia Primatol 1: 188 – 198

Tigges J, Tigges M (1965) Ein Beitrag zur Farb- und Musterbevorzugung bei Gibbons und Orang-Utans. Zool Jb Physiol Bd 71: 337-344

Tomasello M, Call J (1997) Primate cognition. Oxford University Press, Oxford

Visalberghi E, Néel C (2003) Tufted capuchins (*Cebus apella*) use weight and sound to choose between full and empty nuts. Ecol Psychol 15: 215-228

Visalberghi E, Fragaszy D, Ottoni E, Izar P, de Oliveira MG, Andrade FRD (2007) Characteristics of hammer stones and anvils used by wild bearded capuchin monkeys (*Cebus libidinosus*) to crack open palm nuts. Am J Phys 132: 426-444

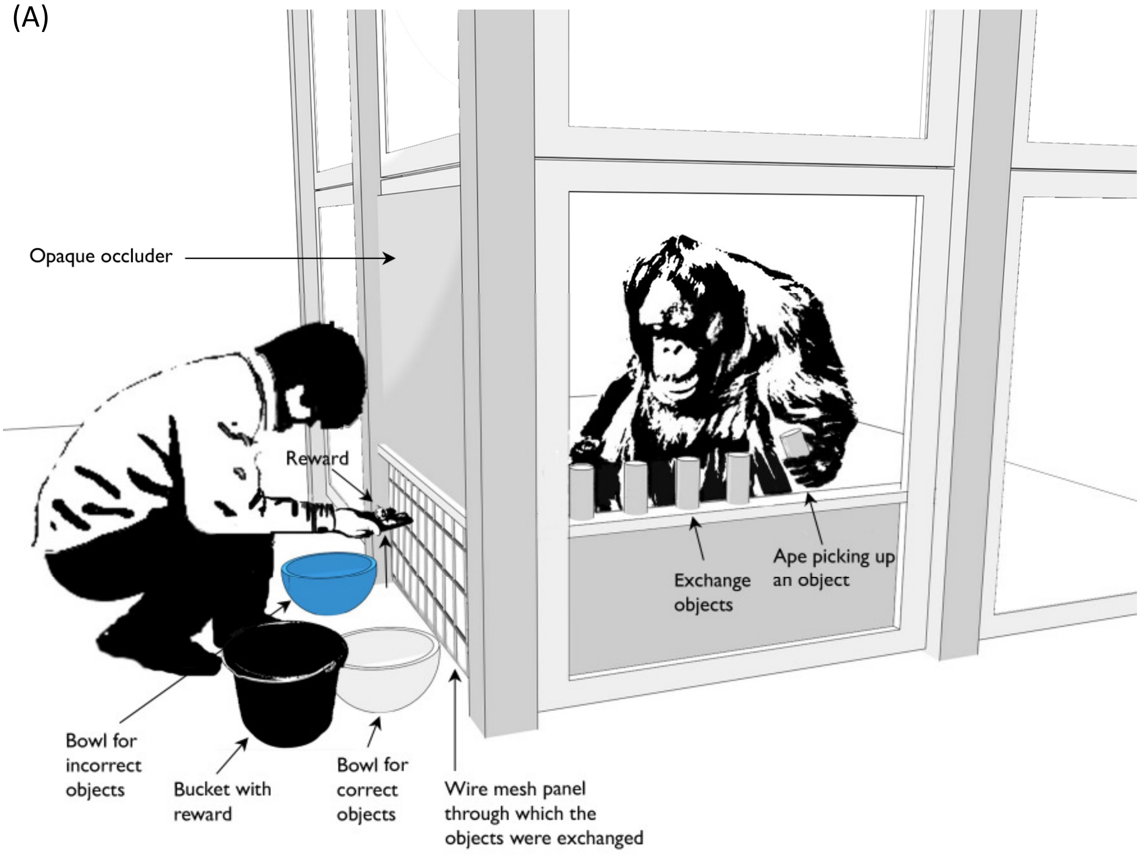
Vlamings PHJM, Uher J, Call J (2006) How the Great apes (*Pan troglodytes*, *Pongo pygmaeus*, *Pan paniscus*, and *Gorilla gorilla*) perform on the reversed contingency task: The effects of food quantity and food visibility. J Exp Psychol Anim Behav Processes 32: 60-70

Vonk J (2003) Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first- and second-order relations. Anim Cog 6: 77-86

Figure 1 Experimental set-up for the exchange experiment: Exchange objects are placed on a ledge inside the ape´s cage. The experimenter shows a reward that will be given to the subject after it has returned a correct object. (A) Ape lifts an object. (B) Ape gives an object to the experimenter. Drawing by Sandra Michaelis.

Table 1 Participants of the exchange experiment. Numbers in bold indicate successful performance. Correct: Type of correct stimulus, Session: Session number in which the criterion (6 out of 6 correct choices) was reached. Order: Order in which the two tests were performed (b-w: Subjects were first tested for color and thereafter for weight, w-b vice versa).

(A)



Opaque occluder
Reward
Exchange objects
Ape picking up an object
Bowl for incorrect objects
Bucket with reward
Bowl for correct objects
Wire mesh panel through which the objects were exchanged

(B)

Table 1: Participants of the exchange experiment.

Species	Subject	Sex	Age	Color Condition		Weight Condition		
				Correct	**Session**	**Correct**	**Session**	**Order**
Orang	Padana	female	8	white	36	light	**30***	b-w
Orang	Pini	female	18	black	36	light	**32***	b-w
Orang	Dokana	female	17	black	**14***	heavy	36	b-w
Orang	Dunja	female	33	white	36	heavy	36	w-b
Orang	Bimbo	male	26	white	36	light	36	w-b
Bonobo	Joey	male	23	black	**29***	heavy	36	b-w
Bonobo	Limbuko	male	10	black	**6***	light	**14***	w-b
Bonobo	Kuno	male	9	white	**3***	light	**9***	w-b
Bonobo	Ulindi	female	12	white	**4***	heavy	**35***	b-w
Bonobo	Yasa	female	9	black	36	heavy	36	w-b
Gorilla	Bebe	female	26	black	**7***	heavy	36	w-b
Gorilla	Viringika	female	11	black	36	heavy	36	b-w

Great apes use weight as a cue to find hidden food

Cornelia Schrauf[1], Josep Call[2]

1 University of Vienna, Department of Cognitive Biology

2 Max Planck Institute for Evolutionary Anthropology, Leipzig

Submitted in:

American Journal of Primatology

Abstract

Bonobos (*Pan paniscus*; n = 5), orangutans (*Pongo pygmaeus abelii*; n = 6) and a gorilla (*Gorilla gorilla gorilla*; n = 1) were presented with two opaque cups, one empty and one baited (containing two bananas). Subjects had to independently gain weight information about the contents of the cups to find the hidden food. Six apes attained above chance level within a total of 16 trials. Successful subjects spontaneously adopted the method of successively lifting the cups and thus comparing their weight before making a choice. Prior to testing, these apes had participated in a weight discrimination task. To rule out that a subject's good performance was influenced by previous experience in weight experiments, we ran a second test in which the same task was presented to a group of chimpanzees (*Pan troglodytes*; n = 9) who were naïve to weight experiments. These subjects also participated in an additional test condition in which the same problem was presented based on learning to associate arbitrary visual stimuli. The results show that experience did not affect performance because, the nine naïve subjects were equally able to find the food when the task stimuli held a causal relation (i.e. perceived weight indicates the hidden food). Interestingly, only one of the naïve subjects solved the task when the task elements held an arbitrary relation (i.e. certain visual pattern indicates food). Our results confirm previous findings that apes perform better in problems grounded on causal compared to arbitrary relations.

Introduction

Contemporary psychology often assumes that animals can learn to associate stimuli and responses, but possess little knowledge about the relation between observed events (e.g. Thorndike 1898; Povinelli 2000; Woodward 2007). Recent studies, however, show that the simple formation of arbitrary connections between certain events and responses is insufficient to explain various phenomena (e.g. Blaisdell et al., 2006; Call, 2004; Seed et al., 2006). Seed & Call (2009) argued that certain species such as apes and corvids perceive functional relations between elements, and do not merely register their co-occurrence. Furthermore, Call (2006) noted that establishing stimuli-response associations is much harder when the elements of a task hold an arbitrary relation rather than a causal/functional relation.

One context in which this discrepancy has been repeatedly observed is when subjects have to use information from an event to infer the location of food (e.g. Call, 2004, 2006; Hanus & Call, 2008). For instance, Call (2004) presented great apes with two opaque containers and gave visual and auditory information about the content of the cups, only one of which contained food. When the cue and the food held a causal relation, i.e. shaking the cup produced either a noise or silence depending on whether there was food inside or not, subjects effectively used the auditory cue to find the food. In contrast, when the cue and the food held an arbitrary relation, i.e. tapping the cup produced the noise, subjects failed to use the cue to find the food. These results are remarkable given that both cues had the same predictive power because they always reliably indicated the presence of food.

Sabbatini & Visalberghi (2008) replicated Call´s experiment with capuchin monkeys (*Cebus apella*), and Schmitt & Fischer (2009) tested olive baboons (*Papio hamadryas anubis*) on the same experimental paradigm. Capuchin monkeys found the food when the provided information derived from shaking the cups (noise or silence produced depending on content). However, the same subjects failed to learn to associate an arbitrary auditory cue with the baited container. Olive baboons were unable to use sound as a cue to find hidden food. Thus, Schmitt & Fischer (2008) restricted the experiment to the use of visual information. First, baboons were shown the content of the empty cup and, based on this information, they had to infer the location of the reward. Six out of 7 subjects solved this problem. In a second experiment, subjects were given an arbitrary cue (a red cross) as an indicator for the location of the reward. Interestingly, only 3 of the 7 subjects learned to correctly use the arbitrary signal. This result is even more compelling because when the arbitrary cue experiment was conducted, the subjects already had pre-experience with the underlying method.

Hanus & Call (2008) also experimentally contrasted task elements that either designated a causal or arbitrary relation. Chimpanzees were confronted with a balance with two opaque cups mounted on either end. After the experimenter placed a piece of banana in one of the cups behind a screen, subjects saw the balance tipping to one side, and chose the lower cup significantly more often than the higher cup (causal condition). In contrast, subjects chose at chance when the experimenter´s action tipped the balance (non-causal balance condition) or when no movement of the balance was involved (non-causal wedge condition). Strikingly, subjects that had been choosing the lower cup in the causal condition reverted to random responses when the weight of the banana was no longer the cause of the movement. This demonstrates that chimpanzees had no intrinsic preference for the lower cup. Again, these findings show that chimpanzees solved the task more easily when the elements held a causal rather than an arbitrary relation.

Two aspects of studies on making inferences about food location remain to be investigated. First, there is the issue of the perceptual modality through which information is presented. In previous studies, subjects witnessed an event that produced either a visual or auditory outcome and then selected one of two options available. It is unclear whether subjects could also use kinaesthetic information such as weight for such a decision. Visalberghi & Néel (2003) tested capuchin monkeys' ability to use sound and weight as cues to infer the fullness of a nut. The authors changed the weight of the nuts by leaving the nutshell either fully rewarded (original state), empty or filled with paper or lead shot. Capuchin monkeys could manipulate both nuts before making a choice. Those subjects touched the nuts and tapped them with their fingers prior to selection, potentially making decisions based not solely on weight, but also on sound and smell. Capuchins' ability to use weight information alone has been investigated in the context of nut-cracking. Visalberghi et al. (2009) found that capuchins can use the weight of stones or artificial hammers (Schrauf et al. 2008) to decide which hammer may be more effective for cracking open nuts. Furthermore, Klüver (1933) showed that these monkeys can also discriminate objects based on their weight, regardless of their functional value, to solve a task such as nut cracking.

Just like capuchins, great apes can also be trained to discriminate reliably between objects of different weight (McCulloch 1941; Schrauf & Call 2009). Mc Culloch (1941) trained five chimpanzees to select the heavier of two boxes. Despite the large weight differences (80 vs. 480g or 80 vs. 640g), the chimpanzees needed a median of 1,100 trials to master the task. Schrauf & Call (2009) found that orangutans and bonobos could also learn to discriminate objects based on their weight. In that study, subjects had to learn to exchange for a reward one of two identical objects that differed only in weight. Some subjects had to exchange the heavier objects, others the lighter objects. Apes mastered the task but required a median of 331 trials to exceed chance-level performance. Subjects also received an equivalent visual discrimination test in which they had to exchange one of two identical objects that differed only in achromatic color (black vs. white). Here, subjects needed a median of 64 trials to reach criterion; this is still a relatively large number of trials, especially considering that, prior to this study, subjects already had experience with the exchange task. Note, however, that McCulloch's (1941) and Schrauf & Call´s (2009) study are both based on solving an arbitrary relation and not a causal one, the correct solution being decided *a priori* by the experimenter.

A second neglected aspect in food location studies is that subjects were never required to seek information themselves about the outcome before making a choice. In all previous studies, subjects were presented with a situation in which they could witness the end result before choosing between the two alternatives, but never produced the outcome themselves. It is therefore important to know whether subjects would play a more active role when the information is not simply provided, i.e. whether they would test and compare different outcomes before making a decision.

Seeking information has been reported in studies devoted to investigating metacognition (e.g. Call & Carpenter, 2001; Hampton et al., 2004; Call 2010). Here, however, the information sought was visual, with on some instances additional auditory information offered (Call 2010), and did not necessarily involve making inferences because subjects could directly spy the food. Visalberghi & Néel (2003) also described capuchins seeking information, but it is unclear whether the information was about weight, sound or perhaps even smell. In the case of nut cracking, Visalberghi et al. (2009) observed capuchins comparing stones for weight and friability but here again, no inference was needed regarding the location of food or any other aspect of the situation.

The current study therefore investigates whether subjects would seek and compare differential outcomes before selecting one of the two alternatives when they had access only to the containers weight. Moreover, we assess whether subjects learn to use weight more easily in a causal situation in which food presence is causally related to cup weight compared to an arbitrary relation as in the Schrauf & Call (2009) exchange study. The rationale for Experiment 1 is to test whether great apes can seek information about weight to infer the food location and whether subjects perform better in this context than in the previous weight discrimination task. Therefore, we presented apes who had already participated in the weight exchange study, with two opaque cups and hid slices of two bananas in one of them. Subjects had the possibility to lift both cups before making a choice. We tested whether apes would independently gain kinaesthetic information about cup weight and subsequently select the heavier one. If subjects learned to use weight to get the reward faster in this task than in the exchange task, this would suggest that subjects privileged the use of causal information.

Experiment 2 was designed to examine if pre-experience in weight discrimination experiments influenced performance in Experiment 1. We also tested whether apes can learn to use an arbitrary cue to locate the food reward when the weight cue is no longer involved and the causal structure thus removed. For this purpose we presented the same task as in Experiment 1 to a group of chimpanzees who were naïve to weight experiments. These subjects also participated in a condition in which cups of the same weight but different visual appearance were presented. The cups were painted with patterns of different color and shape. Subjects had to learn to associate a specific pattern with reward location. Thus, we presented chimpanzees with causal stimuli (i.e. heaviness generated by food inside the cup) and arbitrary stimuli (i.e. a certain visual pattern indicates food), with the contingencies of reinforcement being the same. If apes solved both tasks by establishing arbitrary associations between stimuli (e.g. weight, colors, shapes) and reward, then they should learn the visual cue faster, because previous studies suggest that apes perform better with visual than kinesthetic cues (e.g. McCulloch, 1941; Schrauf & Call, 2009).

Experiment 1

This experiment assessed the general testing procedure and examined whether subjects were capable of using information about weight in order to find hidden food in one of the two cups.

Methods

Subjects

Six orangutans (*Pongo pygmaeus abelii*), five bonobos (*Pan paniscus*) and one gorilla (*Gorilla gorilla gorilla*) housed at the Wolfgang Köhler Primate Research Centre, Leipzig Zoo (Germany), participated in this study. There were 8 females and 4 males ranging from 7 to 34 years of age (Table 1). Apart from one orangutan (Kila), all subjects had been tested on weight discrimination prior to this study (Schrauf & Call 2009) (for a more detailed description, see Table 1). All apes lived in social groups with access to indoor and outdoor compounds. Subjects were tested individually in their indoor cages. The subjects were never food deprived during the experiment and water was available *ad libitum*.

Table 1

Materials

We used two opaque white, conical, paper coffee cups (16.8 cm x 9 cm Ø top side x 6 cm Ø bottom side) with lids (9 cm in diameter) to hide banana slices. The empty and the baited cups weighed 20 g and 320 g, respectively. We attached one end of a string to the bottom side of each cup. A small piece of wood, which served as a stopper, was tied to the loose end of the string. A Plexiglas panel with two oval openings (11.5 cm x 8.5 cm; 40 cm apart) on its upper part served for the presentation of the two cups (Figure 1). A small plastic block (5 x 2.3 x 4 cm) was attached to the Plexiglas above each opening, and the string passed through a slit in the block. The subjects could grasp the string through the opening in the panel. Pulling the string lifted the upside-down cup upwards along the panel and, as soon as it reached the opening, the subjects could touch the cup with their hands. The opening size allowed the apes to grasp the string but prevented them from grabbing the cup. A grey PVC box with an opening at the back (82 x 38 x 30 cm) served to conceal the shuffling of the cups from the subject.

Figure 1

Procedure

The experimenter (E) sat outside the cage and waited until the subject sat in front of the panel to start the trial. First, the E showed the two empty cups to the ape. Then, the E presented a bucket filled with banana slices. E completely filled one of the cups in full view of the subject, whereas the other cup remained empty. The E showed the subject the empty bucket, covered the cups with lids and then shuffled them behind the occluder. Finally, the occluder was removed and the strings were passed through the block on the Plexiglas panel. To indicate their choice, apes had to touch one of the cups with their fingers. To do so, they had to grasp the string and pull it until the cup reached the opening. Only if the subject touched the cup we counted it as choice. Pulling the string and the cup halfway and then releasing it was not counted as a choice. Since holding the string and touching the cup required using both hands, the apes' choices were unambiguous. After a subject made a choice, we unhooked the strings and showed the ape the content of both cups. If the subject had selected the baited cup, it received its content. Otherwise the E put the content of the cup back into the bucket and left the testing room. The location of food was counterbalanced left and right, with the only constraint that it was not hidden in the same location in more than two consecutive trials. Due to the large amount of food (about two bananas per trial), we conducted only 1 trial per day for a total of 16 trials.

Data scoring and analysis

All trials were videotaped. We used two-tailed non-parametric statistics to analyze the data. Our main dependent variable was the number of correct trials (e.g. the number of trials a subject selected the heavy cup). We compared the performance of bonobos and orangutans with the Mann-Whitney test. We used the Wilcoxon test to compare subjects' performance as a group to chance and to analyze whether subjects improved their performance between the first eight trials and the last eight trials. We used the binomial test (expected p=0.5) to analyze the performance on the first trial and whether individual subjects selected the baited cup above chance level. This meant that subjects had to select the correct cup in at least 12 out of 16 trials to be above chance levels. We used Fisher's test to assess the effect that a previous experiment on weight discrimination had on the current study. We also scored whether subjects compared between the two cups by successfully pulling both strings before making a choice. We used Spearman correlation to investigate the relation between success and pulling both strings in each trial. Finally, we used Wilcoxon tests to compare the frequency of pulling the correct and then incorrect cup with the frequency of pulling the incorrect and then the correct cup.

Results

Success

Bonobos performed significantly better than orangutans (Mann-Whitney *U* test *U*=4; N=11; P = 0.05; mean(SEM): bonobos=11.8(SEM=1); orangutans=8.3(SEM=1.6)). All bonobos except one performed above chance levels, whereas only two of the six orangutans did so (Binomial test: P<0.05, Table 2). Subjects as a group did not perform above chance levels (Wilcoxon exact test Z = 1.596; N = 12; P = 0.121). Errors occurred mainly in the first half of the experiment (Trial 1-8) compared to the second half (Trial 9-16) (Wilcoxon exact test Z = 2.223; N = 12; P = 0.03). An analysis of the first trial performance showed no initial preference for the heavy cup (Binomial test P = 0.388). The type of positive stimulus (light/heavy) that had been associated with reward in the Schrauf and Call (2009) weight discrimination study did not designate success in the present study (Fisher exact test P = 0.43). Additionally, success in the former experiment did not determine success in the current one (Fisher exact test P = 0.16).

Table 2

Pulling patterns

Figure 2 presents the percent of successful trials as a function of the number of trials in which subjects successively pulled at both strings. The more subjects pulled on both strings, the more likely they were to succeed in finding the food (Spearman correlation $r(12) = 0.83$, $P = 0.001$). In the first trial, however, only Pini pulled both strings, and ended up selecting the empty cup.

Figure 2

Figure 3 presents the mean number of trials in which subjects performed a particular pulling pattern. Most responses involved a single pull that resulted in a correct or incorrect choice. Moreover, there was no difference in the number of times apes chose the correct or incorrect cup after their first pull (Wilcoxon exact test $Z = 0.359$; $P = 0.773$). However, when subjects switched after their first pull, they did so more often from the incorrect cup to the correct cup (IC) than vice versa (CI) (Wilcoxon exact test $Z = 2.673$; $P = 0.004$). The remaining cases depicted in Figure 3 were too few to enable statistical analyses, but in every case the apes ended up selecting the correct cup.

Figure 3

Discussion

Six of the 12 subjects tested succeeded in using the weight of the cup to find the hidden food. Apparently, the essential step to solve the task consisted in gaining kinaesthetic information about the weight of each of the cups and subsequently selecting the heavier one. Successful apes spontaneously adopted the method of successive comparisons of the cups weight. Compared to the results of Schrauf & Call (2009), subjects showed rapid improvement in learning to discriminate weight even though the weight differences were the same in both studies. This result is even more striking because we used a novel method for the subjects in the present study, whereas in the earlier exchange study subjects were already familiar with the exchange procedure prior to testing. In the current study, subjects might have learned to use weight more easily because of the causal structure of the task, i.e. food causes weight.

Conceivably, the apes´ good performance was influenced by their previous experience with weight experiments, although we detected no significant effect of their previous performance on the current results. Having been tested on weight discrimination prior to the current experiment may have made weight a more relevant dimension to them. Furthermore, one could argue that the fast learning was related to the large amount of food they received as reward. To test these alternative explanations, we ran a second experiment with subjects that had no previous experience on weight discrimination. Additionally, we introduced a second condition that, like the Schrauf & Call (2009) study, was based on an arbitrary (yet equally predictive) relation between the stimuli and the reward. Here, the causal structure was removed as an arbitrary visual cue indicated the location of the reward. If the abundant food promoted the fast learning, and if the task was solved by establishing arbitrary associations between stimuli and the presence of the reward, then the apes should learn the visual cue faster because previous studies suggest that they perform better with visual than weight cues.

Experiment 2

In this experiment we examined the performance of naïve chimpanzees in finding hidden food by using its weight (causal condition), just like in Experiment 1. We also investigated whether they could solve the task by using visual information (i.e. color and shape of stickers) as discriminative cues (arbitrary condition) when such cues were not produced by the presence of food. The reason for adding visual information to weight was because previous studies had suggested greater proficiency in visual compared to weight discriminations.

Methods

Subjects

Nine chimpanzees (*Pan troglodytes*) housed at the Wolfgang Köhler Primate Research Centre, Leipzig Zoo (Germany), participated in this study. The seven females and two males ranged from 6 to 32 years of age (Table 1). All subjects were naïve to weight experiments and were tested individually in their indoor cages. The subjects were never food deprived during the experiment and water was available *ad libitum*.

Materials

We used the same basic apparatus as in Experiment 1 including the two cups, lids and attached strings. For the causal condition we used two white cups, whereas in the arbitrary condition we used two olive green cups which differed in the color (green or blue) and shapes (circles vs. triangles) of the stickers that were attached to them (see Table 3 for information on the precise patterns used). In the causal condition, the empty and the baited cups weighed 20g and 320g, respectively. In the arbitrary condition, both cups weighed the same after the baiting, because the empty cup was filled with a mixture of lead shot and silicone. Therefore, in the arbitrary condition, the weight cue was no longer involved. The cups in the arbitrary condition could only be discriminated by their visual appearance. To hide the lead shot inside the cup from the subject, we glued paper on top of it. Both cups therefore looked identical when they were empty and shown to the subjects. The solid mass formed by the silicone and lead shot distributed the weight evenly in the cup and prevented any rattling noise during manipulations.

Table 3

Procedure

There were two conditions: causal and arbitrary. The order of condition was counterbalanced across subjects. The general procedure was identical to that of Experiment 1. The only changes refer to the arbitrary condition. To make the two conditions as similar as possible, we placed the cups in the arbitrary condition in a grey PVC box (82 x 8.4 x 11 cm) before starting a trial. Without lifting the cups, subjects were unable to see the stimuli that were fixed on the lower half of the cup. Thus, subjects had to lift the cup a bit in order to perceive the stimuli. Each chimpanzee was subjected to 16 trials per condition.

Data scoring and analysis

The same scoring method and analyses as in Experiment 1 was used for the causal and arbitrary condition in Experiment 2. We compared the performance in the causal condition with the arbitrary condition using the Wilcoxon test. This test was also applied to compare subjects' performance as a group to chance and to analyze whether subjects improved their performance between the first eight trials and the last eight trials. We used the binomial test (expected p=0.5) to analyze the performance on the first trial and whether individual subjects selected the baited cup above chance level. This meant that subjects had to select the correct cup in at least 12 out of 16 trials to be above chance levels. We used Fisher's test to assess the effect of order of presentation of condition. To test for species differences, we compared the performance of bonobos, orangutans (Experiment 1) and chimpanzees (Experiment 2 causal condition) using the Kruskal Wallis test. We also scored whether subjects compared between the two cups by successfully pulling both strings before making a choice. We used Spearman correlation to investigate the relation between success and pulling both strings in each trial. Finally, we used Wilcoxon tests to compare the frequency of pulling the correct and then incorrect cup with the frequency of pulling the incorrect and then the correct cup.

Results

Success

Subjects performed significantly better in the causal than arbitrary condition (Wilcoxon exact test Z = 3.728; N = 18; P < 0.001). Moreover, subjects performed above chance level in the causal (Wilcoxon exact test Z = 2.176; N = 9; P < 0.05) but not in the arbitrary condition (Wilcoxon exact test Z = 0.949; N = 9; P = 0.375). Five out of nine chimpanzees performed above chance levels in the causal condition (Binomial test: P<0.05, Table 3), whereas only one did so in the arbitrary condition (Binomial test: P<0.05, Table 3). The order of presentation of conditions had no influence on performance in the causal (Fisher exact test: P=0.318) and arbitrary condition (Fisher exact test: P=0.556).

Comparing the number of errors in the first half (Trial 1- 8) and second half (Trial 9 – 16) of the experiment showed no difference for the causal (Wilcoxon exact test: Z= 1.474; N= 9; P=0.17) or arbitrary condition (Wilcoxon exact test Z=0.954; N=9; P=0.438). In the former, however, only one of the successful subjects (Fifi) made errors in the second half of the study (Trial 15, 16) – and she did so after being correct on 10 successive trials. An analysis of the first trial performance showed no initial preference for either the light or heavy cup (Binomial test: P = 1).

Combining the data from Experiments 1 and 2 (causal condition) showed no significant differences between species (Kruskal-Wallis test: χ^2= 3.97; N= 20; P = 0.137; mean(SEM):bonobos=11.8(SEM=1),orangutans=8.3(SEM=1.6), chimpanzees=11.7(SEM=1.2).

Pulling patterns

Figure 4a & b show the percent of successful trials as a function of the number of trials in which both strings were pulled successively. The more both strings were pulled, the greater the success in finding the food both in the causal (Spearman correlation r(9)=0.755, P=0.019) and arbitrary conditions (Spearman correlation r(9)=0.777, P=0.014). In the causal condition, 4 subjects compared the cups on their first trial, 3 of them ending up with the correct cup. In the arbitrary condition, 2 subjects compared the cups on their first trial, both ending up with the correct cup.

Figure 4 a & b

Figure 5a presents the mean number of trials in which subjects performed a particular pulling pattern in the causal condition. In those cases involving a single pull (C & I), there was no difference in the number of times apes chose the correct or incorrect cup (Wilcoxon exact test Z = 1.577; P = 0.125). However, when subjects switched after their first pull, they did so more often from the incorrect to the correct cup (IC) than vice versa (CI) (Wilcoxon exact test Z = 2.388; P = 0.016). With three exceptions, on all remaining occasions in which subjects compared the cups more often (two to five times), they always ended up selecting the correct cup.

Figure 5a

Figure 5b shows the mean number of trials in which a certain pulling pattern was performed in the arbitrary condition. In single pulls (C & I), apes chose the incorrect cup more often than the correct cup (Wilcoxon exact test Z = 2.673; P = 0.004). When they switched after their first pull, they did so more often from the incorrect to the correct cup (IC) than vice versa (CI) (Wilcoxon exact test Z = 2.226; P = 0.031). In the remaining cases in which subjects switched between cups, they ultimately chose the correct cup.

Figure 5b

Discussion

Five of the nine chimpanzees tested were able to find the food in the causal condition (i.e. food inside the cup generated perceived weight). In contrast, only one subject solved the task in the arbitrary condition, where visual cues indicated the correct cup. Interestingly, in the causal condition naïve subjects showed a higher frequency of multiple comparisons per trial before making a choice than experienced subjects in Exp 1. Moreover, whereas experienced subjects showed a maximum of three comparisons per trial, naïve subjects compared the cups four and even five times per trial. The effect was opposite for the arbitrary condition. In particular, multiple comparisons per trial occurred less often in the arbitrary versus causal condition as well as compared to Experiment 1. These findings show that all apes react to weight cues, both those with experience in weight discrimination and naïve ones. However, naïve subjects were less sensitive to weight differences and thus compared more than experienced subjects. Because in the arbitrary condition more comparisons per trial do not enhance the probability to locate the hidden food, it is reasonable why apes compared less often in that condition.

Subjects´ good performance in the causal condition corroborated our findings from Experiment 1 and clearly showed that success in the causal condition cannot be explained by previous experience on weight experiments. Furthermore, when the weight cue and thus the causal structure was removed, subjects were unable to select the correct cup above chance levels even after 16 trials. Despite the visual cues, only one subject was able to solve the task in the arbitrary condition. This makes it unlikely that abundant food accounts for the fast speed of learning in the causal condition of both experiments.

General Discussion

Great apes were able to find hidden food in one of two opaque cups using information about the food´s weight. Success in this task required seeking kinaesthetic information about the weight of each of the cups and subsequently selecting the heavier one. Apes spontaneously adopted the method of successively pulling the strings and thus comparing weights before choosing. Unlike what we observed in Schrauf & Call (2009), subjects showed rapid improvement in learning to discriminate weight although the weight differences were the same in both studies. Whereas subjects needed a median of 331 trials until mastery in the weight discrimination task (Schrauf & Call 2009), here they reached above chance level within 16 trials. This result is even more remarkable because, when the exchange task in the earlier study was conducted, the subjects already had pre-experience with the underlying method, the only difference being the tested feature (weight). In contrast, the current study involved a method that was novel for the subjects.

We presented task elements that held a logical relation, (i.e. heaviness generated by food inside the cup) and task elements with no logical connection between the stimuli (i.e. a certain visual pattern indicates food). Only one subject was able to solve the problem when the task elements involved arbitrary relations. Subjects that began with the causal condition and had successfully solved the task reverted back to chance level in the arbitrary condition. This result is congruent with Hanus and Call´s (2008) findings on the balance beam problem. In their study, chimpanzees that had chosen the lower cup in the causal condition (when food caused the movement) reverted back to chance when the causal structure was removed and arbitrary stimuli served as cue for the location of food.

The difference in performance (causal versus arbitrary condition) and the speed of learning between the causal condition and the earlier weight discrimination experiment (Schrauf & Call 2009) can be explained by the structure of the presented tasks. The aspect of arbitrary and nonarbitrary relations between the cue and the food plays a major role in performance. In the causal condition of the current study, the presence of food is causally related to cup weight. Thus, weight is an important factor contributing to success, as the perceived weight indicates the hidden food. Subjects witness that a large amount of bananas is hidden in one of two cups and have to link the presence of food with its weight. Here, weight is a reliable cue because the cup with the bananas will always be heavier than the empty one. In contrast, in the arbitrary condition the causal structure is removed. There is no obvious reason for a subject why a certain visual stimuli should be correct because it is simply decided by the experimenter. In this arbitrary context the visual cue is unreliable a priori because there is no logical relation between the stimuli. Thus, subjects have to learn to associate a certain visual stimulus with the food location. The same effect of arbitrariness is evident in the earlier weight discrimination task.

We can rule out the possibility that performance differences between conditions are due to weight being a more salient cue than color or shape. As Schrauf & Call (2009) showed, subjects learned color discrimination faster than weight discrimination. And, if one of the conditions had a more salient cue, then it would have been the arbitrary condition because individuals could rely on a certain color and/or shape. Furthermore, we can exclude that the abundant reward in the present study promoted learning: subjects in the arbitrary condition received the same amount of food, and still only one subject was able to solve the task. Finally, we can rule out the possibility that fast learning reflects previous experience in weight experiments because we ran the same experiment with a group of naïve chimpanzees. Even though these subjects had not been tested on weight experiments in the past, they also performed above chance level within a total of 16 trials. Note also that half of the participants in the weight conditional discrimination task had learned to respond to light weight as the type of positive stimulus.

Neither subjects with weight learning experience nor naïve subjects showed an initial preference for the heavy or light cup in this task. After a few trials, successful apes introduced the method of successively pulling the strings before making a choice. Note that we did not present the weights consecutively, but the apes themselves found a way to compare the cups in sequence. The essential step to success seems to involve seeking information about the weight of both cups by successively comparing them before choosing. These findings are consistent with Klüver's (1957) weight discrimination results with Java monkeys using the pulling-in technique with moveable boxes. In particular, three Java monkeys were presented with two boxes of unequal weight and had to respond to the heavier box to get rewarded. Subjects had to pull the strings that were attached to the boxes to bring them within reach. Monkeys showed no improvement in learning for a median of 310 trials until the first appearance of the comparison behavior, i.e. the successive displacement of the boxes. Thereafter the errors decreased rapidly (p. 28).

Our successful subjects started to compare the cups already after the first trial (Median = Trial 2). The apes determined the sequence of comparison in a given trial as well as the number of comparisons. After their first pull, they switched significantly more often from the incorrect to the correct cup than vice versa. Naïve subjects (Exp 2), however, showed a higher frequency of multiple comparisons per trial, before making a choice, than experienced subjects (Exp 1). Moreover, whereas experienced subjects showed a maximum of three comparisons per trial, naïve subjects compared the cups four and even five times per trial. A possible explanation for this difference is that experienced subjects had an advantage by being already more sensitive to the weight differences. When Klüver decreased the weight differences in his experiments from 450 g:150 g to 325 g:250 g, the number of comparisons subjects did per trial increased. He explained these findings by arguing that "the great numbers of comparisons were undoubtedly indicative of the great uncertainty of the monkeys (p. 62)". This explanation is further supported by the fact that, in the arbitrary condition, multiple comparisons per trial occurred less often compared to the causal condition in Experiment 2 as well as compared to Experiment 1. In the arbitrary condition, more comparisons per trial did not enhance the probability to locate the hidden food, which may explain why apes compared less often in that condition.

In summary, apes quickly learned to use weight information in a situation in which weight held a causal relation with the task solution. Subjects failed to solve the same problem when it was based on learning to associate arbitrary visual stimuli. Our results confirm previous findings (Call 2004; 2006a, b; 2007; Hanus and Call 2008) that apes perform better in problems grounded on causal compared to arbitrary relations.

References

Blaisdell AP, Sawa K, Leising KJ, Waldmann MR. 2006. Causal reasoning in rats. Science 311: 1020-1022.

Call J. 2004. Inferences about the location of food in the Great Apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla,* and *Pongo pygmaeus*). J Comp Psychol 115: 159-171.

Call J. 2006a. Descartes´ two errors: Reason and reflection in the great apes. In: Hurley S, Nudds M, editors. Rational animals. Oxford: Oxford University Press. p 219-234.

Call J. 2006b. Inferences by exclusion in the great apes: the effect of age and species. Anim Cog 9: 393-403.

Call J. 2007. Apes know that hidden objects can affect the orientation of other objects. Cognition 105: 1-25.

Call J. 2010. Do apes know that they could be wrong? Anim Cog DOI 10.1007/s10071-010-0317-x

Call J, Carpenter M. 2001. Do apes and children know what they have seen? Anim Cog: 3, 207-220.

Hampton RR, Zivin A, Murray EA. 2004. Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. Anim Cog 7: 239–254.

Hanus D, Call J. 2008. Chimpanzees infer the location of a reward based on the effect of its weight. Curr Biol 18: R370-R372.

Klüver H. 1933. Behavior mechanisms in monkeys. Chicago University Press, Chicago.

McCulloch TL. 1941. Discrimination of lifted weights by chimpanzees. J Comp Psychol 32: 507-519.

Povinelli DJ. 2000. Folk Physics for Apes: The Chimpanzee´s Theory of How the World Works. Oxford University Press, Oxford.

Sabbatini G, Visalberghi E. 2008. Inferences about the location of food in capuchin monkeys (*Cebus apella*) in two sensory modalities. J Comp Psychol 122: 156-166.

Schmitt V, Fischer J. 2009. Inferential Reasoning and Modality Dependent Discrimination Learning in Olive Baboons (*Papio hamadryas anubis*). J Comp Psychol 123: 316-325.

Schrauf C, Huber L, Visalberghi E. 2008. Capuchin monkeys learn to use weight to select the most effective tool to crack open nuts. Anim Cog 11: 413-422.

Schrauf C, Call J. 2009. Great apes´ performance in discriminating weight and achromatic color. Anim Cog 12: 567-574.

Seed AM, Tebbich S, Emery NJ, Clayton NS. 2006. Investigating Physical cognition in Rooks, *Corvus frugilegus*. Curr Biol 16: 697-701.

Seed AM, Call J. 2009. Causal knowledge for events and objects in animals. In: Watanabe S, Blaisdell AP, Young A, editors. Rational Animals, Irrational Humans. Keio University Press, p 173-187.

Thorndike EL. 1898. Animal intelligence. An experimental study of the associative processes in animals. Psychol Rev Monogr Suppl 2: 551-553.

Visalberghi E, Néel C. 2003. Tufted capuchins (*Cebus apella*) use weight and sound to choose between full and empty nuts. Ecol Psychol 15:215–228

Visalberghi E, Addessi E, Truppa V, Spagnoletti N, Ottoni E, Izar P, Fragaszy D. 2009. Selection of effective stone tools by wild capuchin monkeys. Curr Biol 19: 213-217.

Woodward J. 2007. Interventionist theories of causation in psychological perspective. In: Gopnik A, Schultz L, editors. Causal Learning: Psychology, Philosophy and Computation. Oxford University Press, Oxford p 16-36.

Table 1 Species, name, age, sex, previous experience in weight experiments and experiment participation of the subjects

Table 2 Participants of Experiment 1. Numbers in bold indicate successful performance. Weight exchange correct: type of correct stimulus in the exchange experiment. Weight exchange performance: Successful: subjects reached required criterion. Failure: subjects did not meet criterion.

Table 3 Participants, pairing of presented cups and correct stimulus of Experiment 2. Numerals represent in how many trials subjects were correct in the respective condition. Numbers in bold indicate successful performance. Note that a minimum of 12/16 is necessary to perform above chance level. Order: order in which the two tests were performed (c-a: subjects were first tested in the causal and thereafter in the arbitrary condition, a-c: vice versa)

Figure 1 Experimental set-up of the testing apparatus: Subject pulls the string to lift the cup towards the opening. The cups were visually identical but differed in weight. Drawing by Sylvio Tuepke

Figure 2 Number of comparisons over the course of the experiment in relation to success for all subjects in Experiment 1. Primary axis: Number of trials in which subjects successively pulled on both strings. All pulls are considered, independent of whether subjects ended up with the correct or incorrect cup. Secondary axis: Percent of correct trials. Dashed line indicates above chance performance ($P < 0.05$)

Figure 3 Mean number of times subjects showed the following behavior: C= correct I= incorrect C or I = single pull; CI or IC = 1 comparison; CIC = 2 comparisons; ICIC = 3 comparisons; First letter indicates the cup subjects started to pull; Last letter indicates the cup subjects ended up with

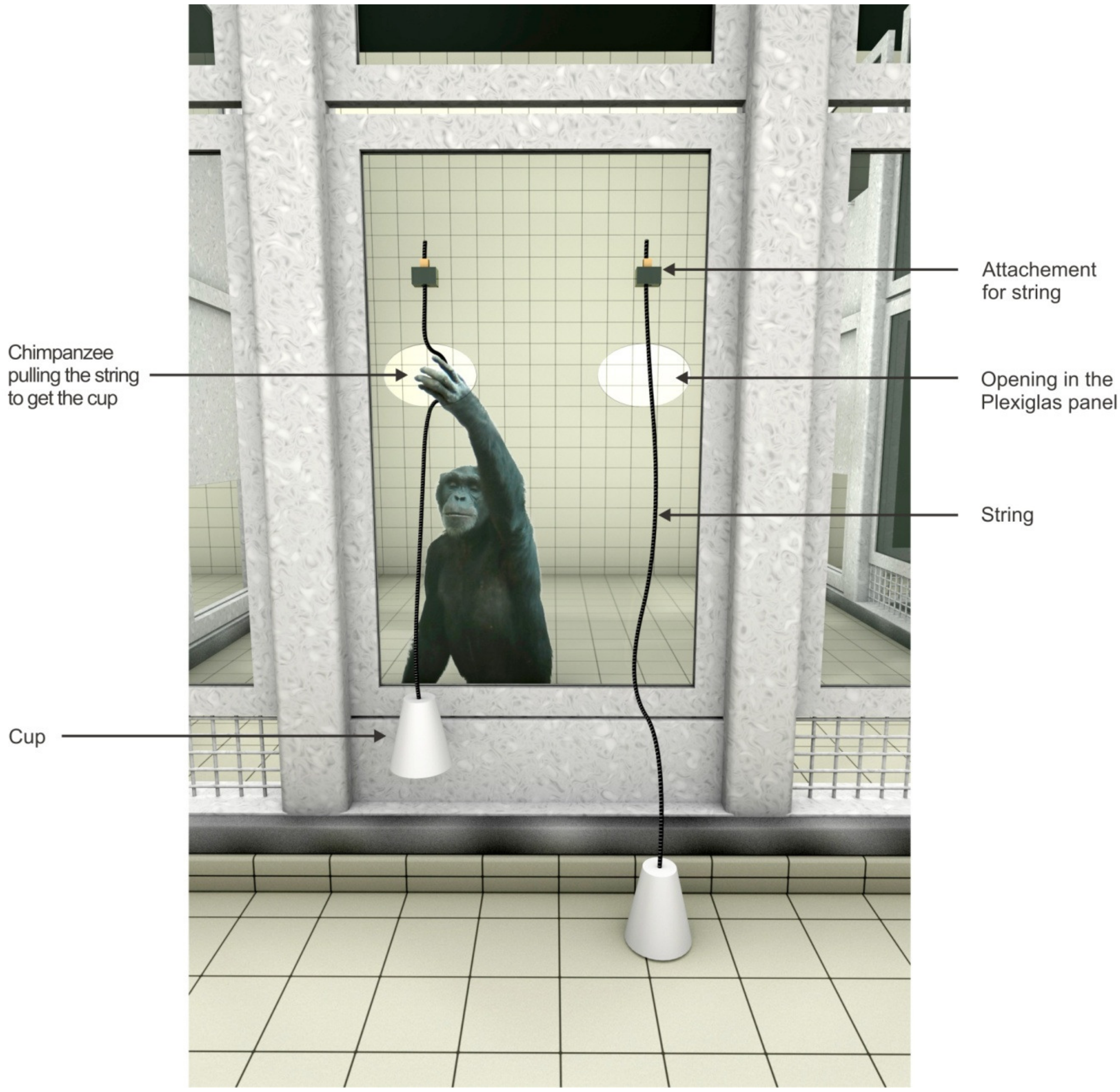
Figure 4 A causal condition; B arbitrary condition. Number of comparisons over the course of the experiment in relation to success for all subjects in Experiment 2. Primary axis: Number of trials in which subjects successively pulled on both strings. All pulls are considered, independent of whether subjects ended up with the correct or incorrect cup. Secondary axis: Percent of correct trials. Dashed line indicates above chance performance ($P < 0.05$)

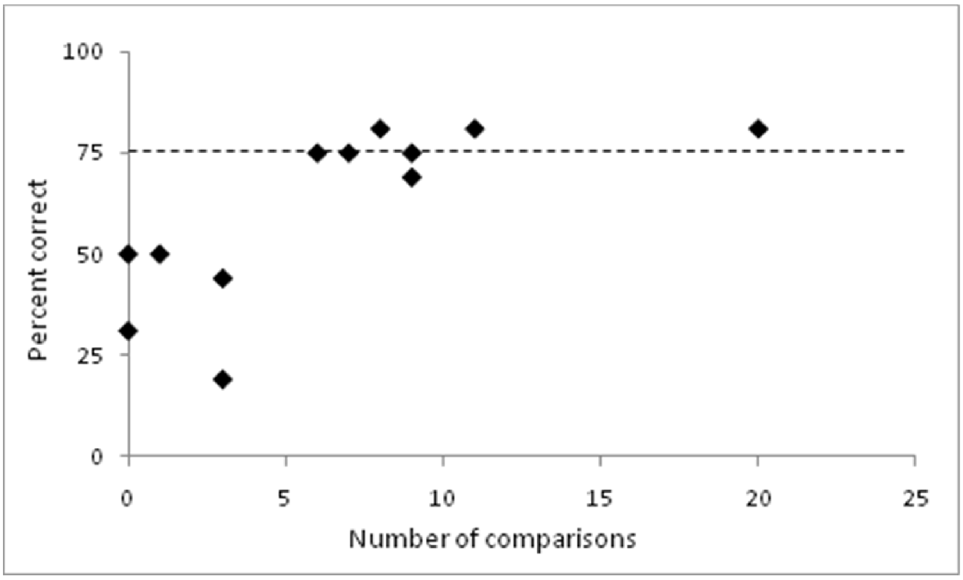
Figure 5 A causal condition; B arbitrary condition. Mean number of times subjects showed the following behavior: C= correct I= incorrect C or I = single pull; CI or IC = 1 comparison; CIC = 2 comparisons; ICIC = 3 comparisons; CICIC= 4 comparisons; CICICI= 5 comparisons. First letter indicates the cup subjects started to pull; Last letter indicates the cup subjects ended up with

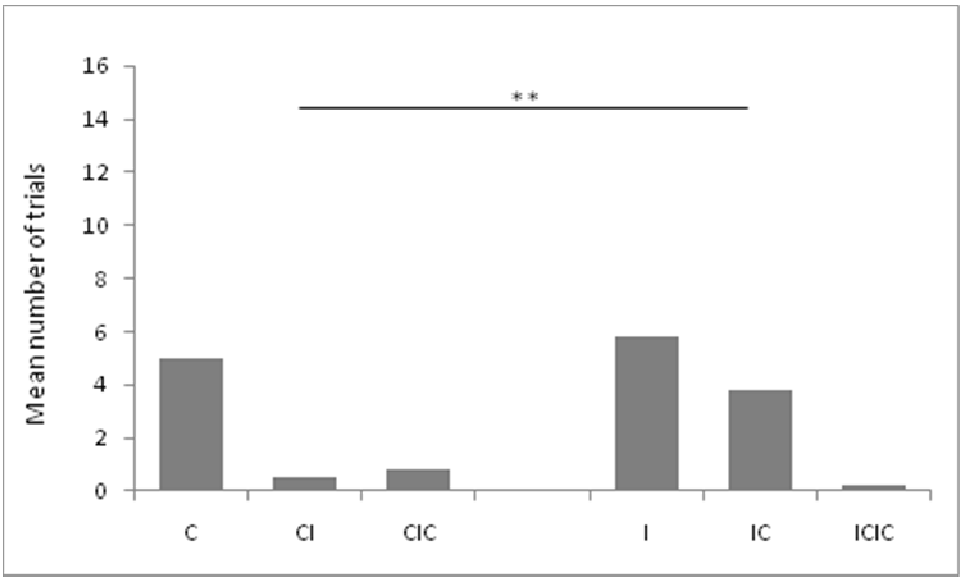
Species	Subject	Age (years)	Sex	Previous experience in weight experiments	Experiment participation
Orangutan	Padana	10	F	pass	1
Orangutan	Dokana	18	F	fail	1
Orangutan	Pini	19	F	pass	1
Orangutan	Dunja	34	F	fail	1
Orangutan	Bimbo	27	M	fail	1
Orangutan	Kila	7	F	not tested	1
Bonobo	Joey	25	M	fail	1
Bonobo	Kuno	11	M	pass	1
Bonobo	Limbuko	12	M	pass	1
Bonobo	Ulindi	14	F	pass	1
Bonobo	Yasa	10	F	fail	1
Gorilla	Viringika	12	F	fail	1
Chimpanzee	Alex	7	M	not tested	2
Chimpanzee	Annette	9	F	not tested	2
Chimpanzee	Alexandra	9	F	not tested	2
Chimpanzee	Jahaga	15	F	not tested	2
Chimpanzee	Gertruida	15	F	not tested	2
Chimpanzee	Fifi	15	F	not tested	2
Chimpanzee	Pia	8	F	not tested	2
Chimpanzee	Fraukje	32	F	not tested	2
Chimpanzee	Lome	6	M	not tested	2

Species	Subject	Performance	Weight exchange	
			Correct stimulus	Performance
Orangutan	Padana	**12**	Light	Pass
Orangutan	Pini	7	Light	Pass
Orangutan	Dokana	**12**	Heavy	Fail
Orangutan	Dunja	11	Heavy	Fail
Orangutan	Bimbo	5	Light	Fail
Orangutan	Kila	3		Not tested
Bonobo	Joey	8	Heavy	Fail
Bonobo	Limbuko	**13**	Light	Pass
Bonobo	Kuno	**12**	Light	Pass
Bonobo	Yasa	**13**	Heavy	Fail
Bonobo	Ulindi	**13**	Heavy	Pass
Gorilla	Viringika	8	Heavy	Fail

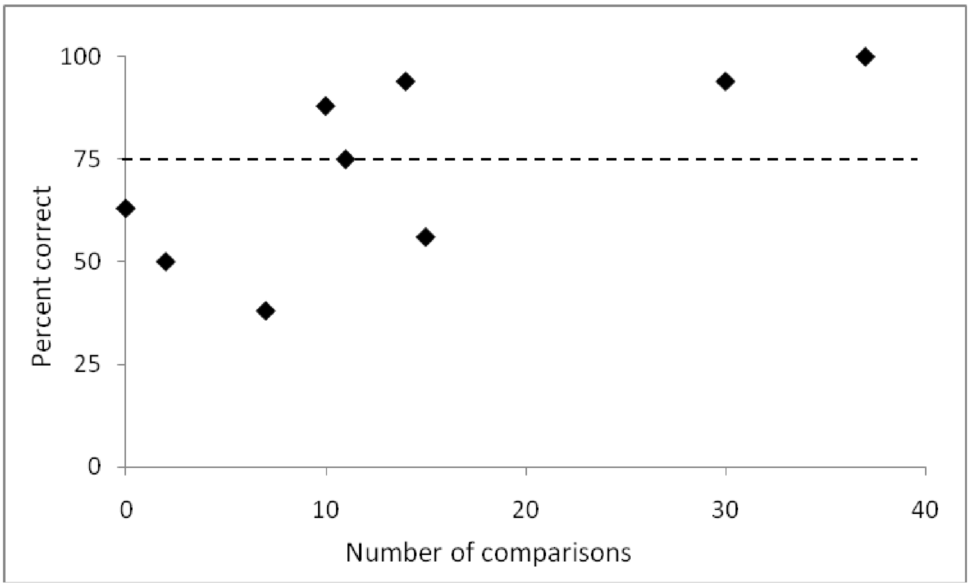
Subject	Pair of presented cups	Correct stimulus	Condition		Order
			Arbitrary condition	**Causal condition**	
Pia	green O vs. yellow Δ	green O	8	9	c-a
Fraukje	blue Δ vs. green O	blue Δ	7	10	c-a
Lome	blue Δ vs. green O	blue Δ	8	8	c-a
Jahaga	blue Δ vs. green O	blue Δ	5	**16**	a-c
Alexandra	yellow Δ vs. blue O	yellow Δ	7	**14**	a-c
Fifi	yellow Δ vs. blue O	yellow Δ	6	**12**	c-a
Anette	green O vs. blue Δ	green O	6	**15**	a-c
Gertruida	green O vs. blue Δ	green O	8	6	a-c
Alex	blue Δ vs. green O	blue Δ	**13**	**15**	c-a
		Mean	7.6	11.7	
		% correct	48	73	


Attachement
for string
Chimpanzee
pulling the string
to get the cup
Opening in the
Plexiglas panel
String
Cup

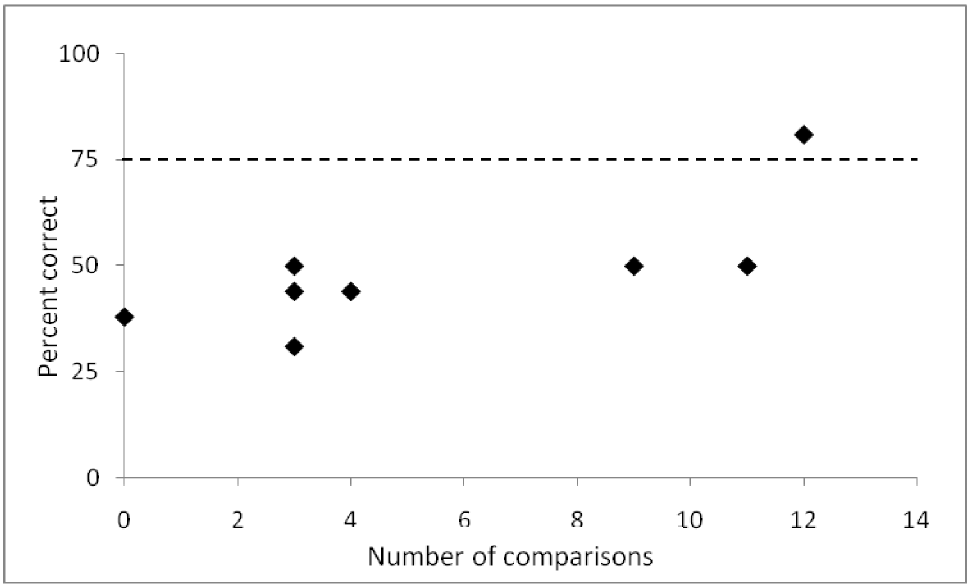

100
75
50
25
0
Percent correct
0
5
10
15
20
25
Number of comparisons


16
14
12
10
8
6
4
2
0
Mean number of trials
**
C
CI
CIC
I
IC
ICIC

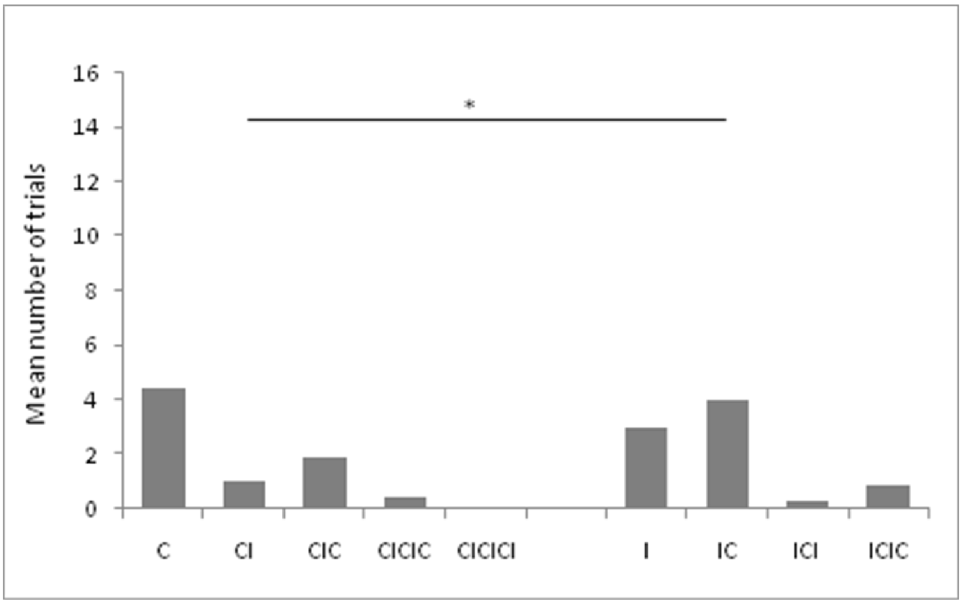
(A)


100
75
50
25
0
Percent correct
0
10
20
30
40
Number of comparisons

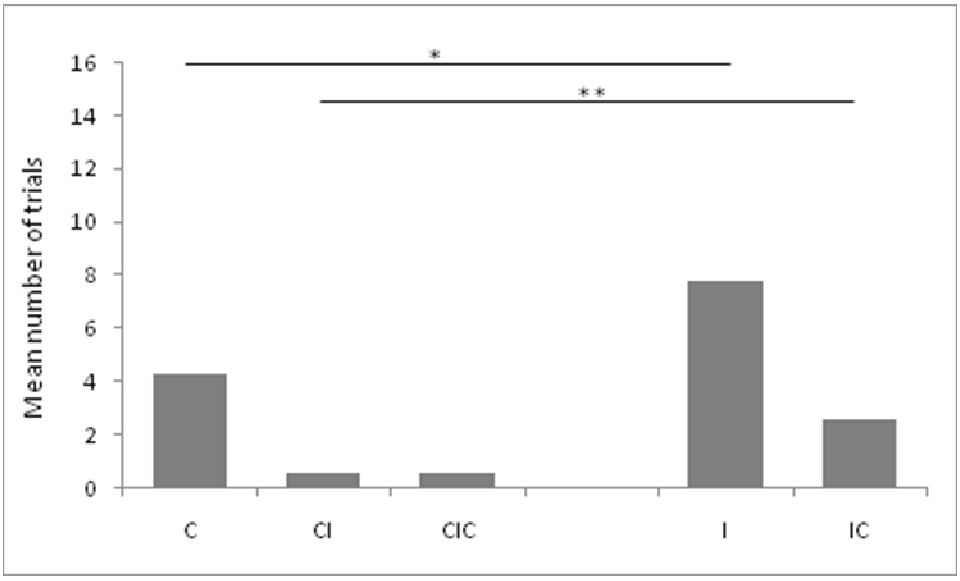
(B)


100
75
50
25
0
Percent correct
0
2
4
6
8
10
12
14
Number of comparisons

(A)


16
14
12
10
8
6
4
2
0
Mean number of trials
*
C
CI
CIC
CICIC
CICICI
I
IC
ICI
ICIC

(B)


16
14
12
10
8
6
4
2
0
Mean number of trials
*
**
C
CI
CIC
I
IC

The effect of physically possible versus impossible balance-scale feedback on the expectancies of 3- to 4-year-old children

Cornelia Schrauf[1], Josep Call[2], Sabina Pauen[3]

1 University of Vienna, Department of Cognitive Biology

2 Max Planck Institute for Evolutionary Anthropology, Leipzig

3 University of Heidelberg, Institute for Psychology

Submitted in:

Journal of Cognition and Development

Abstract

Previous studies (Siegler, 1981; Case, 1985) have shown that children under the age of 5-years have little understanding of balance-scales when required to integrate information about both weight and distance from the fulcrum. However, an understanding based on weight alone is assumed to be present even in 2-year-olds (Halford *et al*., 2002). In all of these studies, weight was varied using multiple objects of the same weight. As a result, the children's decisions could have been based upon visual features (size, number) without necessarily taking weight into account. The present study investigated whether young children are able to correctly encode the relevance of weight in influencing the behavior of a balance-scale. We studied how well 3- to 4-year-old children learn to use one of two different weights (of equal appearance) to tip the scale. In one condition (possible), the heavy weight produced the desired outcome, in the other condition (impossible), the light weight caused the scale to tip. Only 4-year-olds learned more effectively in the possible versus impossible condition, suggesting that children younger than 4 have not yet developed clear expectations of the role of weights on the movements of a balance-scale.

Introduction

Children in all cultures have some knowledge about the physical world, generally long before they receive any formal education in physics. Intuitive physics is an area of research that studies our beliefs about physical events occurring in everyday life, predating any systematic theory of physics. Numerous studies have found an early competency in infants' physical beliefs, suggesting that humans are innately endowed with representational knowledge about physical phenomena (e.g. Carey, 1985; Gelman, 1991; Keil, 1989; Leslie, 1995; Spelke *et al*., 1992). In these experiments, infants are presented with a series of events – some that obey the laws of physics (such as solidity or gravity) and others that violate these laws. Several studies have shown that human infants respond differently (e.g. by looking longer) when there is a violation as opposed to a conformity with the physical laws (e.g. Baillargeon, 1995; Spelke *et al*., 1995) - even though they cannot translate their thinking into a coherent verbal account of their performance. Despite a growing body of evidence for an intuitive physics, however, our intuitions about physical causation are not always correct. There are also some deeply held intuitive misconceptions about physical phenomena, even in human adults (McCloskey, 1983).

The balance-scale task belongs to the most frequently used task paradigms in developmental psychology to investigate certain aspects of children's understanding of the physical world (Wilkening & Huber, 2002; Pauen & Wilkening, 1997; Krist *et al*., 2004). Originally introduced by Inhelder and Piaget (1958), the balance-scale task consists of a bar placed on top of a fulcrum or balance point. Weights can be positioned on both sides of the fulcrum, and the effectiveness of a weight in causing the scale to tip is determined by a combination of its weight and its distance from the fulcrum. The classic task has several variations not only involving different perceptual demands but also varying in terms of the types of required output. More specific, in the absence of feedback and proprioceptive information, children had to predict what would happen, if a balance-scale with a given arrangement of weights was free to move. In another variation of the task, children were allowed to act themselves in adjusting the objects´ distance from the fulcrum to achieve equilibrium of the balance. In the 1970s and early 1980s the balance-scale task was used extensively to construct and investigate theories about learning and cognitive development in children (e.g. Inhelder & Piaget, 1958; Koslowski & Bruner, 1972; Siegler, 1976; 1978; 1981; Case, 1985). Scientists were particularly interested in children´s understanding of the way objects balance and whether they are able to integrate information from two different dimensions (weight and distance). Apart from that, the balance-scale task is also an assignment in which children often balance objects with considerable success, but cannot explain how they achieved this success (Karmiloff-Smith & Inhelder, 1974).

Findings to date (e.g. Case, 1985; Inhelder & Piaget, 1958; Siegler, 1981; Surber & Gzesh, 1984) indicate that children have little understanding of the balance-scale before 5 years of age. Case (1985) stated that children up to the age of 5 years would only be able to decide which side of the balance would go down when there was a gross perceptual difference between the two stacks of weights. Karmiloff-Smith and Inhelder (1974) stated that it is not until the age of 7 that children start to consider weight as a significant property in problems involving equilibrium; only later on are children able to differentiate between weight as an absolute property and weight as a force. In contrast, Halford *et al*. (2002) claimed that an initial understanding of the balance-scale based on weight alone is present even in 2-year-olds.

Recently, Messer *et al*. (2008) argued that the discrepancy in results across different balance tasks may be related to children not utilizing the same cognitive processes across different tasks. In particular, performance appeared to be greatly influenced by perceptual and task characteristics.

Many of the previously described studies assumed that children´s initial representation of the elements on the beam is framed in terms of weight (e.g. Case, 1985; Siegler, 1978). However, in the experimental procedure on which these conclusions are based, an unequal number of identical elements is placed on each end of the scale to achieve differences in weight, rather than presenting visually identical heavy and light objects. Thus, number prediction pattern is confounded with weight prediction pattern. Metz (1993) utilized an experimental procedure in which the weight of the elements varied. More specifically, she presented 3-to-5-year-old children with a pan balance and a set of several weights that were either identical or differed in size and/or weight. Children had to distribute the weights across the pans to achieve equilibrium. In case the placement of weight resulted in disequilibrium, children could use this feedback to fix the problem. Fifteen of the 48 subjects tested never thought of the elements as weight, including nine 3-year-olds and six 4-year-olds. These children first placed the same number of elements in each pan. When by doing so they failed to achieve equilibrium (in some cases one element contained hidden extra weight), the children tried to fix the apparatus instead of adjusting the placements of the weights. In particular, subjects carefully placed the apparatus in the correct alignment and then gently released or yanked the apparatus into the correct position. Based on these findings the author concluded that younger children may not represent the elements on the beam in terms of weight but in terms of quantity in a collection. In contrast, all but two 5-year-olds always assumed the relevance of weight to the problem of the task.

Children as young as 2 years of age can discriminate objects by their weight (Robinson, 1964). It remains open, however, whether younger children are able to understand the relevance of weight to the balance-scale problem. If children possess intuitive beliefs about basic lever principles and understand how weight enters into task resolution, they should have an implicit expectancy about the role of weight when it comes to predicting the behavior of a balance-scale. Hence, children´s expectations should affect learning performance by reinforcing the learning of relations that are consistent with their prior experiences and by hindering the learning of relations that violate these expectations and contravene the laws of physics.

Following this line of reasoning, the present study investigates whether young children (i.e. 3- to 4-year-olds) were able to correctly encode the relevance of weight in influencing the behavior of a balance-scale. More specifically, we studied how well children learn to use one of two given weights (of similar appearance) to tip a scale. We presented only two objects, from which children had to select one, to eliminate confounding of numerical, non-weight or weight representation.

In the causal (possible) condition, only the heavy object produced the desired effect of disrupting the equilibrium of the scale, while in the inverse (impossible) condition, only the light object was effective. In the latter case the movements of the scale were controlled by the experimenter beyond the participant's view. The inverse condition was used to test whether children would learn an arbitrary association between weight and the tilting of the balance as quickly as a causal one. If children possess a prior intuitive knowledge about weight which can be used to predict the movement of a balance-scale, they would be expected to perform better in the causal than in the inverse condition.

Method

Participants

Sixty children participated in this study. These children were split into three age groups (3, 3.5 and 4 years old), with every child's birthday falling within a range of ± 14 days of their respective age class. There were 20 children in each age group, with 10 participants per condition. Half of the children were assigned to the causal condition and the other half were assigned to the inverse condition. The number of boys and girls was equal in each group. All participants were recruited from a list of parents who had volunteered to participate in child development studies. The children were tested in a child laboratory, which they visited with a parent for one testing session of approximately 20 minutes. None of the participants had any experience with a balance-scale prior to the experiment.

Materials

Apparatus

The balance scale consisted of two parallel bars (76.5 cm x 3.5 cm, distance between them: 5.3 cm) and two fulcrums (39 cm x 18 cm and 39 cm x 24 cm, respectively, see Fig. 1). The scale was placed on top of a grey PVC box (85.5 cm x 55.5 cm x 42.6 cm) and presented to the child behind a Plexiglas panel (81.4 cm x 55 cm x 32.8 cm) to prevent children from yanking or aligning the scale with their hands. In the upper half of the panel were two openings (7 cm in diameter) through which the children could reach the balance-scale. On both ends of the scale's arm, we mounted a metal dish (12 cm x 7 cm). Both dishes were covered with black felt so that any objects placed on top would not slide off. Whenever the scale was in an inclined position (irrespective of side) the uppermost dish was aligned exactly with one of the openings, thus allowing the child to reach through and grab the object from the dish.

The grey PVC box had an opening at the back, and two small holes were drilled through the board resting on top of the PVC box (see Fig. 2). We attached a transparent string to both arms of the balance-scale and inserted the other ends through the holes in the PVC board. Finally, we attached metal rings to the ends of the strings: the experimenter could now control the movements of the balance-scale by pulling the rings with her hands.

The experimenter manipulated the strings beyond the child's view, sitting behind the apparatus, whereas the children were told to remain in front of the apparatus during the experiment. The children's designated area was indicated by a boundary-line painted in red on the floor.

Figure 1

Figure 2

Stimulus materials

We used two rectangular, yellow boxes (6 cm x 5 cm x 4 cm) of identical outward appearance as stimulus material. One of the boxes was filled with lead shot, weighing a total of 320 g. The other box remained empty, weighing only 20 g. A transparent silicon paste was added to the lead shot to bind the contents into a solid and homogeneous mass, preventing any rattling noise during manipulations and evenly distributing the weight throughout the tool. A stuffed mouse (*Lillebi*) with a pink shoulder bag was used as an incentive for the participants. The mouse's bag contained either a stamp or a small piece of cookie that could be collected by the children as a reward.

Procedure

After a warm-up phase the children were accompanied to the experimental room by their parent and two experimenters. Inside, the parent was asked to take a seat at a distance from the apparatus and was instructed to remain quiet during the testing session. Children were tested individually and were assigned to one of two groups: Group 1 received the causal condition (i.e. the heavy object tilted the scale) and Group 2 participated in the inverse condition (i.e. the light object tilted the scale). In the inverse condition, the experimenter prevented the tipping of the scale whenever a heavy object was used and produced a downward movement of the balance whenever a light object was applied.

All subjects were presented with a total of 12 trials. To avoid fatigue or loss of motivation, a short break was introduced after the first six trials. During the break the children played with an apparatus that had proven to be an effective reward in previous experiments (Warneken & Tomasello, 2008). Here, when the child threw a cube through an opening in the apparatus, the cube slid down a transparent tube into a box and created a pleasant jingling sound.

Demonstration

Children received several demonstrations to better understand the mechanisms of a balance-scale. In addition, children themselves manipulated the apparatus by exerting pressure with their hands to tilt the scale. In particular, one experimenter (henceforth, E1) presented the balance-scale to the child. E1 and the child walked together to the back of the apparatus, where E1 pulled down the upper arm of the scale with her hand and then pointed with her finger to the opposite side saying: “Look when this side goes down, the other side goes up.” This was done for both sides. Then, in a second phase, the child was asked to do the same. The aim of this demonstration was twofold: First, we wanted to show the children that pressing down one arm of the scale raises the opposite arm and, because both arms of the scale are connected, they will not move independently of one another. Second, we wanted to give the children the opportunity to feel for themselves the resistance of the balance when pressing down one side with their hands. These proprioceptive explorations fostered children´s understanding of the balance-scale.

For the second demonstration, E1 walked with the child to the front of the apparatus. The child was told to watch what Experimenter 2 (E2) was doing. E2 walked behind the balance-scale and placed an object on one of the scale’s mounted dishes, leaving the opposite dish empty. This caused the balance to tip so that the side bearing the object touched the table. The object was then removed, the equilibrium was restored and the same procedure was applied using the other dish. This demonstration was designed to show that, in the same manner as ones hand can press down the scale’s arm and thereby uplift the opposite side, the placement of elements can pull down one side and raise the opposed side.

The last demonstration included the stuffed mouse that would later participate in the actual experiment. E1 directed the child’s attention to the balance-scale. E2 placed the mouse on one dish and an object on the other. This action caused the mouse to move upwards into a position directly behind the opening in the Plexiglas barrier. The child was then allowed to take the mouse. This demonstration gave the child an opportunity to feel the weight of the mouse (57 g). Furthermore we wanted the child to experience retrieving the mouse from the dish by reaching through the opening in the Plexiglas barrier after the scale had been tilted.

Experiment

The child and E1 remained in front of the apparatus. E1 said "*Now we are going to play a game*". E2 (behind the apparatus) showed the child the stuffed mouse with the pink bag. E1 asked the child "*Do you see this mouse? She has just been on a journey and has returned with a present for you. The present is in her bag*". E2 positioned the mouse on one of the dishes. This tilted the scale, with the empty dish coming to rest in the upper position directly behind the opening in the Plexiglas barrier. E1 then showed the child two stimulus objects saying "*I'm going to give you two objects. One in one hand and the other in the other hand*". Then the objects were placed in the child's hands. "*You can place one of the objects on the dish to get the mouse. But only one of them*". E1 pointed to the opening where the empty dish was resting in the upwards position. The child was then required to place one of its objects on the empty dish (see Fig. 1).

If the child's choices were correct, then the mouse carrying the reward in her shoulder bag moved upwards: the child could then take the mouse from the dish, open its bag and obtain the reward. Placing an incorrect object on the empty dish produced no movement, and E1 said "*Hmmh.. do you want to try it again?*"

In each trial, the position of the heavy object (in the child's left or right hand), as well as the position of the mouse (on the left or right dish) was semi-randomized in the sense that the same side was not chosen more than twice in a row, and both sides were chosen equally often over all 12 trials.

Verbal responses

After the test session, E1 took the two stimulus objects and placed one in each of the child´s hands. Then the child was asked "*Are the boxes the same or are they different?*" If the child did not reply, E1 asked "*When you hold them in your hands do they feel the same? Why?*"

Data analysis

All sessions were videotaped and scored from the footage. We coded the identity of the objects the children placed on the dishes. Our main dependent measure was the number of errors the children made during the experiment.

Because the number of errors followed approximately a Poisson distribution, we calculated a Poisson regression (function 'glm' with family=poisson, in R, R Development Core Team, 2007) with condition (causal, inverse), age (3-, 3.5- and 4-year-olds) and sex as independent variables. We tested the significance of the interaction using a likelihood-ratio test comparing the deviance of the full model (including the interaction) with that of a reduced model (excluding the interaction). We checked for model-validity by visual inspection of the residuals plotted against the predicted values. This indicated no severe deviation from the assumption of constant variance.

Based on the number of errors for each child, we also determined whether a given child was successful in solving the presented problem. We considered those children to be successful that chose correctly in at least 10 out of 12 trials (Binomial test: $P<0.05$). To test for a possible difference between children's initial state versus final state, a Wilcoxon-exact-test was used to compare the first versus second half of trials.

Finally, we looked at children´s first trial performance to test for an initial preference for either the light or heavy object by using the Binomial test.

Results

Error analysis

The median number of errors as a function of condition and age is depicted in Fig. 3. A test of the interaction between condition, age and sex revealed a significant effect on the children's performance (Likelihood-ratio test: $Z(1) = 3.695$; $P < 0.001$). There was no main effect for age $Z(1) = -1.658$; $P = 0.097$, condition $Z(1) = -0.280$; $P = 0.780$ and sex $Z(1) = 1.381$; $P = 0.167$. The two-way interaction for age and condition $Z(1)= 0.493$; $P = 0.622$) and age and sex $Z(1)=-1.330$; $P = 0.184$ revealed no significant effect, but there was an interaction between condition and sex $Z(1)=-3.776$; $P < 0.001$.

As demonstrated in Fig. 3, the number of errors in the causal condition decreased as a function of age, whereas the opposite was true for the inverse condition. More precisely, in the causal condition, 4-year-old children made far fewer errors than 3- and 3.5-year-old children. In the inverse condition, however, 4-year-old children made more errors than the children in the other two age groups.

Figure 3

Table 1 reports the number of errors, separated for condition and age. All 4-year-old children who received the causal condition reached the criterion of 10/12 correct choices. In the inverse condition, only half of the children from the 4-year-old group reached the criterion. In contrast, 3- and 3.5-year-old children showed no differences in the percentage of successful subjects between conditions. An analysis of first trial performance showed no initial preference for either the light or heavy object for all age groups (3-year-olds: Binomial test P= 1; 3.5-year-olds: Binomial test P= 0.503; 4-year-olds: Binomial test P=0.824). Comparing the number of errors in the first half (Trial 1-6) and the second half (Trial 7-12) of the experiment revealed no difference for both conditions in the 3-year-olds (Causal condition: Wilcoxon-exact test Z=-0.141, P=0.984; Inverse condition: Wilcoxon-exact test Z=-0.086, P=1). The same was true for 3.5-year-olds (Causal condition: Wilcoxon-exact test Z=-1.638, P=0.120; Inverse condition: Wilcoxon-exact test Z=-2.236, P=0.063). In 4-year-olds, errors occurred mainly in the first half of the experiment compared to the second half (Causal condition: Wilcoxon-exact test Z=-2.646, P=0.016; Inverse condition: Wilcoxon-exact test Z=-2.333, P=0.031).

With regard to gender differences, in the causal condition boys and girls performed similarly across all age groups. However, this was not the case for the inverse condition: 4-year-old girls made far more errors than boys of the same age group. From a total of five 4-year-old girls, two made 12/12 errors and only one reached criterion. In contrast, 4 out of 5 boys in the same age class were successful. Interestingly, in the 3.5-year-old category, boys made more errors than girls. One boy made 12/12 errors and, whereas all 3.5-year-old girls reached criterion, only two boys did so. No sex differences in performance were found in the group of 3-year-old children. Boys and girls made errors to almost the same extent.

Table 1

Verbal responses

Six 3-year-old children responded to E1's questioning but none of them invoked the term "*heavy*" or "*light*". In contrast, all ten 4-year-old children who answered did invoke this distinction (e.g. 'this one is heavy'). In the 3.5-year-old category, 7 out of 11 children referred to one of the objects as heavy (or light) and 5 of those had reached the criterion of 10/12 correct. In general, most children who responded verbally were the successful ones. The number of boys and girls who spoke was equal for all age classes.

Discussion

The children's performance on the balance-scale task was influenced by age, condition and gender. We found that 3- and 3.5-year-old children performed similarly in the inverse and the causal condition, whereas 4-year-olds performed better in the causal than in the inverse condition. In fact, 4-year-olds performed better than younger children in the causal condition, but not in the inverse condition. Moreover, 4-year-old boys performed better than 4-year-old girls in the inverse condition, whereas the reverse was true for 3.5-year-olds.

A key finding was that, with increasing age, performance improved in the causal condition but decreased in the inverse condition. The behavior of one 3.5-year-old boy and two 4-year-old girls in the inverse condition was particularly striking because they persevered in using the heavy object to tilt the balance in 12 out of 12 trials. These children did not switch to the lighter object in a single trial even though the heavy object continually failed to produce the desired outcome. This behavior strongly indicates that children assume imbalance had to be caused by the object's weight.

Independent of age, children showed no initial preference for the heavy or light object. Only 4-year-old children showed significantly more errors in the first versus second half of the experiment for both conditions. In the causal condition, 4-year-old children´s performance was unambiguous. All participants reached criterion with at most a single error in the first trials. In contrast, in the inverse condition we found a discrepancy in the 4-year-old's performance. They either solved the task without exceeding a single error in the first trials, or they showed no improvement over the course of the experiment, with two children producing even 12 out of 12 errors.

This discrepancy implies that the ability to link the relevance of weight to the balance-scale task is not fully manifested even at this age class. Indeed, some 4-year-olds seemed to fail to understand the influence of weight on the behavior of a balance-scale, indicated by their willingness to learn a rule that contravenes the laws of physics, namely that the light object tilts the scale.

That younger children performed equally well in both conditions (possible and impossible condition) might suggest that these subjects were more likely to attain the goal state (i.e., tilting of the scale) by focusing on the visual feedback of the apparatus then understanding the role of weight here. More specifically, children were sensitive to the feedback they received from the balance-scale and relied on this feedback in terms of whether the last placement had tilted the scale. Because visual feedback is iterative and constitutes an effective route to success in getting the reward, the children could easily apply this strategy. In these cases the weight of the elements functions as an object identifier (perceived heaviness) on a non-causal basis in the same way as color. Thus, children consider the unequal weights of the elements but without understanding how exactly it enters into task resolution. In contrast, older children go beyond responding to the apparatus feedback because their performance differs between conditions. They apparently extend their representation of felt weight to the effect of differential weights placed on the apparatus.

Thus, whereas younger children seem to understand weight as felt weight (i.e. felt pressure while holding an object), older children considered weight as a physical property of objects that causally affects that object's interaction with the balance. This is indicated by their strikingly different performance when confronted with a situation that violates their initial expectations, namely that a heavy object will tilt the scale.

Our results confirm those of Krist *et al*. (2004), who presented 4-year-old children with either a possible event (i.e. the scale tipped down on the side bearing the larger number of objects) or an impossible one (i.e. the scale tipped down on the side bearing the smaller number of objects). Similar to the current study, Krist *et al*. (2004) found more learners in the condition with the possible (42%) than the impossible event (24%). However, more children solved the task in the current study than in Krist *et al*.'s (2004) study – both in the condition with the possible (causal) event (100%) and the condition with the impossible (inverse) event (50%). Moreover, Krist *et al*. (2004) found that 4-year-olds needed several trials to master the task, whereas in our study this age group did it very quickly. One possible reason for these discrepancies is that Krist *et al*. (2004) varied the number of objects placed on each end of the scale rather than presenting visually identically heavy and light objects, as we did. As we used two boxes of similar appearance but different weight, children had to rely solely on their perception of the object's weight when deciding which box should be placed on the scale. This difference in experimental design is important because it is crucial to distinguish the context of numerical, non-weight representation from weight representation. Several previous studies have differentiated weight in terms of the quantity of presented objects rather than weight per se (e.g. Case, 1985; Siegler, 1978). Thus, children's decisions could therefore potentially be based on visual features (size, number) without needing to consider weight.

Metz (1993) addressed this issue in investigating 3-5-year-old children's developing knowledge of the pan balance. Children were presented with a balance and a set of several weights that were either identical or differed in size and/or weight. The task consisted in distributing the weights across the pans to achieve equilibrium. Younger children (3-4-year-olds) achieved this by putting an equal number of elements in each pan. By relying on the apparatus feedback after their placements, they achieved the goal state. However, when this strategy failed (when one element contained hidden extra weight), instead of adjusting the weight placements, children tried to fix the apparatus by yanking and aligning it with their hands. Older children, in contrast, accounted for such inequality (e.g. hidden extra weight) by the idea of trade-offs between the number of elements in a pan and their respective weights. Three-year-olds never compensated between weight and number of discrete elements when displacing the objects across the pans to modify the beam alignment. Metz concluded that younger children encode the contents of each pan in terms of number of elements and their relative weights. Only 5-year-olds and some of the 4-year-old children showed a fully elaborated weight-based approach to the task, representing the elements as weights and recognizing the relevance of weight to the task.

In our study, most 4-year-olds and those 3.5-year-olds who passed the possible (causal) condition invoked the heavy-light distinction in their verbal responses after the test. Three-year-old children never used the term "heavy" when asked about the differences between both weights. It is unlikely that this youngest group generally lack the term "heavy". In fact, their parents reported that the children were well aware of this term and used it, for example in situations when they lifted large books for their parents to read to them, or when they helped carry the groceries. One plausible explanation is that although the term 'heavy' might be employed by children in certain situations, they may not apply it to others, such as in the current balance-scale task. In other words, these children might appeal to the heaviness of an object to explain some manner in which the object affects themselves, but not as a physical property of the object itself.

Smith *et al*. (1985) probed 3-to-9-year-old children's understanding of the words "larger" and "heavier" on a series of paired objects. Every age group made errors by saying "same", although the two cylinders were always of different weight. When the objects were the same size but differed in weight, especially the 3-year-olds (44%) said "same weight". Only 11% of the 4-year-olds did so.

All the above findings indicate that, with increasing age, children acquire experience on the influence of weight on basic lever principles. This implicit knowledge affects learning performance in a balance-scale task by reinforcing the learning of relations that are consistent with the children's prior experiences (causal condition) and by hindering the learning of relations that are inconsistent with them (inverse condition). As long as children of a given age-range have not yet developed any rule-based knowledge to guide their expectations of what occurs under normal circumstances (i.e. circumstances consistent with the causal condition), they seem to be perfectly willing to learn a rule contradicting physical causality (see inverse condition). This is not the case, however, when they have already started to form initial beliefs. Then, they stick to their newly developed ideas and remain fairly resistant for some time to contradicting empirical evidence (Karmiloff-Smith & Inhelder, 1974).

In our study differences between the causal and inverse condition were not found in 3-year-olds, but began to be seen in some 3.5-year-olds and started to become manifested in the 4-year-olds. Some of the older children in our sample possibly already possessed intuitive beliefs about basic lever principles and understood how weight enters into task resolution. Younger children might not have an implicit expectancy of the relevance of weight in predicting the behavior of a balance-scale. This is indicated by their willingness to learn a rule that is not in accordance with the laws of physics. Instead, they seemed to rely more on the visual feedback of the apparatus as a route to success.

In conclusion, our study showed age-related differences in 3-4-year-old children's understanding of weight as a cause for the movement of a balance-scale. Whereas younger children did not differ in their performance between causal (possible) and inverse (impossible) conditions, 4-year-olds performed better in the former. Our results suggest that younger children have not yet developed clear expectations regarding the role of weight in basic lever principles.

References

Baillargeon, R., Kotovsky, L., & Needham A (1995) The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack & A.J. Premack (Eds). *Causal cognition - A multidisciplinary debate*. Oxford: Clarendon Press.

Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, MA: Bradford Books, MIT Press

Case, R. (1985) *Intellectual development: Birth to adulthood*. New York: Academic

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp 293-322). Hillsdale, NJ: Erlbaum Associates

Halford, G.S., Andrews, G., Dalton, C., Boag, C. & Zielinski, T. (2002). Young children's performance on the Balance Scale: The Influence of Relational Complexity. *Journal of Experimental Child Psychology,* 81, 417-445

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*. New York: Basic Books

Karmiloff-Smith, A., & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition,* 3, 195-212

Keil, F.C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press

Koslowski, B., & Bruner, J. S. (1972). Learning to use a lever. *Child Development*, 43, 790–799

Krist, H., Bach, S., Öndül, S. & Huber, S. (2004). Mikrogenetische Studien zum physikalischen Wissenserwerb von Kindern: Neue Trainingsexperimente mit der Balkenwaage. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie,* 36, 119-129

Leslie, A.M. (1995). A theory of agency. In D. Sperber, D. Premack & A.J. Premack (Eds.) *Causal cognition: A multidisciplinary debate* (pp 121-149). Oxford, UK Clarendon Press

McCloskey, M. (1983). Intuitive physics. *Scientific American,* 248, 122-130

Messer, D.J., Pine, K.J. & Butler, C. (2008). Children's behaviour and cognitions across different balance tasks. *Learning and Instruction,* 18, 42-53

Metz, K.E. (1993). Preschoolers´ Developing Knowledge of the Pan Balance: From New Representation to Transformed Problem Solving. *Cognition and Instruction,* 11(1), 31- 93

Pauen, S. & Wilkening, F. (1997). Children's analogical reasoning about natural phenomena. *Journal of Experimental Child Psychology*, 67, 90-113

Robinson, H.B. (1964). An experimental examination of the size weight illusion in young children. *Child Development*, 35:91–107

Siegler, R.S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, 8, 481-520

Siegler, R.S. (1978). The origins of scientific reasoning. In R.S. Siegler (Ed.) *Children´s thinking: What develops?* (pp 109-149). Hillsdale, NJ: Erlbaum

Siegler, R.S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, 46, 1-84

Spelke, E.S., Breinlinger, K., Macomber, J. & Jacobson, K. (1992) Origins of knowledge. *Psychological Review*, 99, 605-632

Spelke, E.S., Philips, A.T. & Woodward, A.L. (1995). Infants knowledge of object motion and human action. In D. Sperber, D. Premack, A. Premack (Eds) *Causal cognition: A multidisciplinary debate*. Oxford, UK: Oxford University Press

Surber, C.F. & Gzesh, S.M. (1984). Reversible operations in the balance scale task. *Journal of Experimental Child Psychology*, 38, 254-274

Warneken, F. & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds*. Developmental Psychology*, 44(6), 1785 - 1788.

Wilkening, F. & Huber, S. (2002). Children's intuitive physics. In Goswami U. (Ed.) *Handbook of childhood cognitive development* (pp.349-370). Oxford: Blackwell

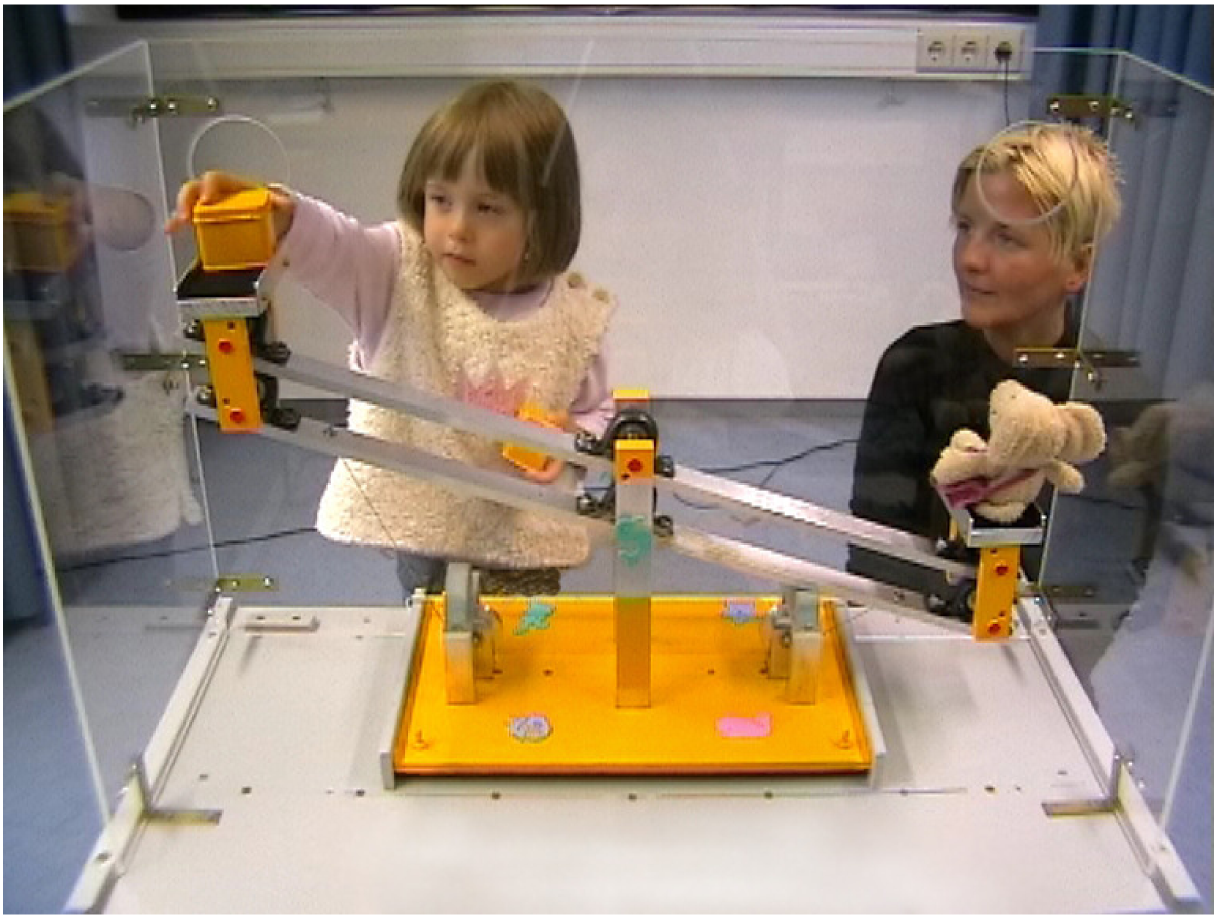
Table 1 Median number of errors on the balance-scale task separated for age and condition. Percentage of children who succeeded in the task separated for age and condition.

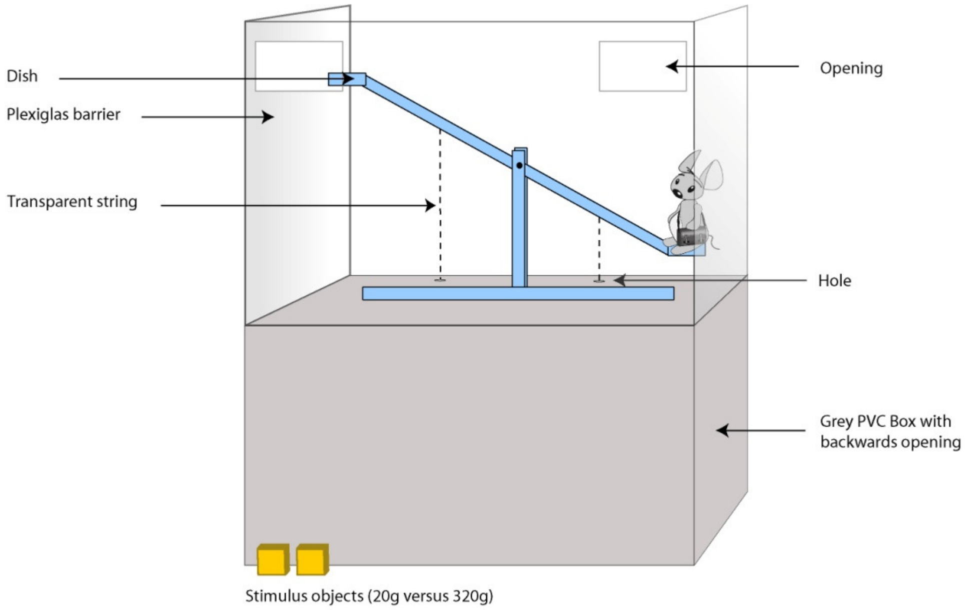
Fig. 1. Child places one of the two given objects on the dish of the balance-scale.

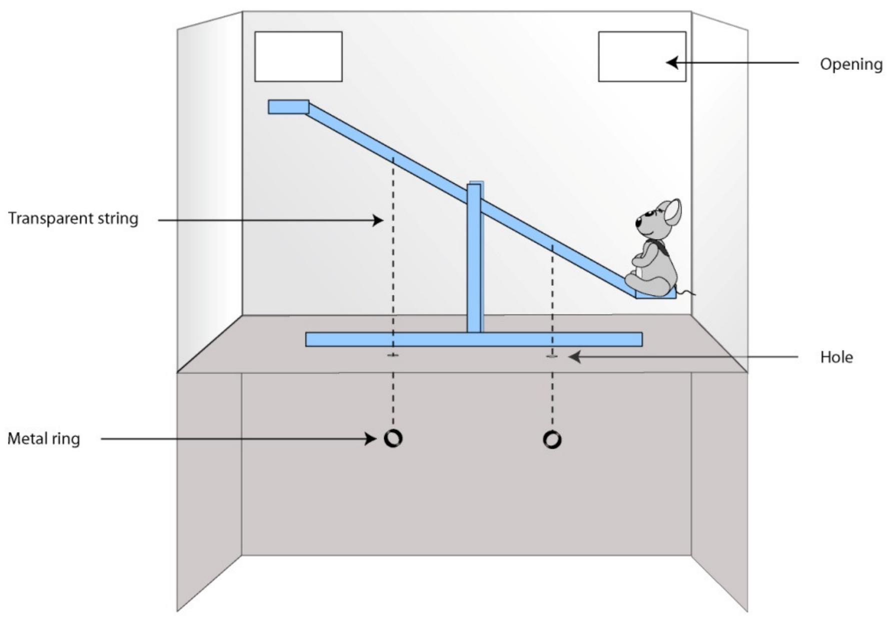
Fig. 2. Experimental set-up of the balance-scale task.

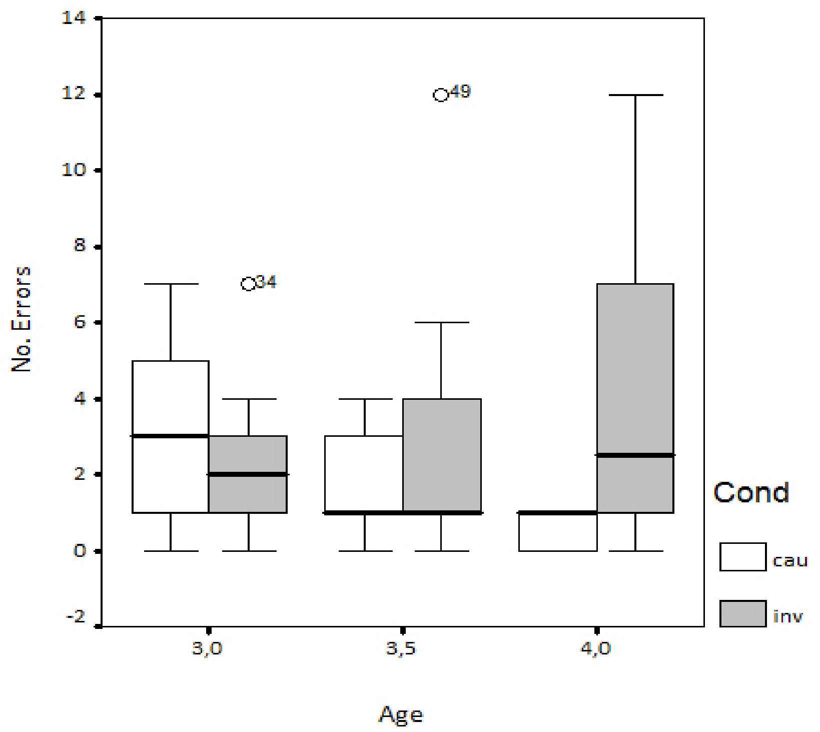
Fig. 3. Median number of errors in 3-, 3.5- and 4-year-old children by condition

		Median # Errors (Range)		% Successful children	
Age	**N**	**Causal**	**Inverse**	**Causal**	**Inverse**
3	20	3 (0-7)	2 (0-7)	50%	70%
3.5	20	1 (0-4)	1 (0-12)	60%	70%
4	20	1 (0-1)	2.5 (0-12)	100%	50%




Dish
Plexiglas barrier
Transparent string
Opening
Hole
Grey PVC Box with backwards opening
Stimulus objects (20g versus 320g)


Opening
Transparent string
Hole
Metal ring



14
12
10
8
6
4
2
0
-2
No. Errors
O34
O49
3,0
3,5
4,0
Age
Cond
cau
inv

Zusammenfassung

Im Laufe des letzten Jahrhunderts ist auf dem Gebiet der visuellen Fähigkeiten von Primaten viel Forschungsarbeit geleistet worden, von einer physiologischen (Kremers 2005) und kognitiven Perspektive (Tomasello & Call 1997). Im Gegensatz dazu ist über die Relevanz von nicht-visuellen Informationen bei Primaten wenig bekannt. Eine Vielzahl von Pflanzen und Früchten ändern ihre Farbe und andere visuelle Eigenschaften während des Reifungsprozesses nur wenig wenn überhaupt. In Anbetracht dessen, sollte die Verwendung anderer Sinne auch eine wichtige Rolle spielen um relevante Informationen zur Nahrungsmittelauswahl zu erlangen. Povinelli (2004) vertritt die Ansicht, dass Schimpansen auf sichtbare Eigenschaften von Objekten angewiesen sind und kein Verständnis von nicht sichtbaren Objekteigenschaften wie zum Beispiel Gewicht haben. Nach der Meinung des Autors ist die menschliche Sprache eine wesentliche Vorbedingung um die Wahrnehmung von Gewicht mit einem Konzept von Gewicht zu verbinden. Im Rahmen meiner Doktorarbeit habe ich untersucht, ob Menschenaffen und 3-4 jährige Kinder ein Verständnis von Gewicht als nichtvisuelle Objekteigenschaft besitzen. Desweiteren bin ich der Frage nachgegangen, ob sie in der Lage sind anhand von Gewichtsinformationen Schlussfolgerungen zu ziehen. In der ersten Studie wurde untersucht, ob andere Menschenaffen als Schimpansen eine Aufgabe zur Gewichtsdiskrimination lösen können. In der Folgestudie wurden alle Menschenaffenarten getestet. Konkret habe ich geprüft, ob Menschenaffen Gewichtsinformation in einer Situation verwenden können, in welcher Gewicht in einer kausalen Beziehung zur Lösung der Aufgabe steht. Abschließend wollte ich untersuchen, ob 3-4 jährige Kinder in der Lage sind die Relevanz von Gewicht als physikalische Kraft an einer Balkenwaage zu erkennen. Die Ergebnisse zeigen, dass Menschenaffen Gewichtsunterschiede von Objekten wahrnehmen können und Gewicht als Information zum Lösen einer kausalen Problemstellung verwenden können. Menschliche Kinder unter 4 Jahren haben noch keine klaren Erwartungen hinsichtlich der Rolle von Gewicht auf das Verhalten einer Balkenwaage entwickelt. Diese Resultate legen nahe, dass die menschliche Sprache nicht die Vorbedingung für ein abstraktes Verständnis von Gewicht ist.

Abstract

Over the last century much work has been done on the visual abilities of primates, both from a physiological (Kremers 2005) and cognitive perspective (Tomasello & Call 1997). In contrast, very little is known about the relevance of non-visual information in primates. Given that a variety of plants and fruits change color or other visual features only little if at all during maturation, other senses might also play an important role in acquiring relevant information for food selection. Povinelli (2004) suggested that chimpanzees rely strictly upon observable features and have no understanding of unobservable entities such as weight. In the authors opinion human language is essential to connect perception of weight with a concept of weight. My PhD centered on the study of great apes´ (orangutans, bonobos, gorillas and chimpanzees) and 3- to -4-year-old children's understanding of weight as a non- visual- based object property and whether they are genuinely able to reason about weight. In the first experiment I evaluated how great apes, other than chimpanzees, perform in a weight discrimination task. In a next step I investigated whether these apes and a group of naive chimpanzees would be able to use weight information in a situation in which weight has a causal relation with the solution of the task. Finally, I wanted to assess whether 3-to-4-year-old children are able to correctly encode the relevance of weight as a physical force in influencing the behavior of a balance- scale. The results showed that great apes are sensitive to differences in weight and can use weight information in a problem-solving task. Human children younger than 4 years have not yet developed clear expectations regarding the role of weights on the movements of a balance-scale. These findings suggest that human language is not mandatory for an abstract understanding of weight.

Curriculum Vitae

Mag.rer.nat. Cornelia SCHRAUF

PhD student

Department of Cognitive Biology

University of Vienna

connyschrauf@chello.at

Personal Information

Nationality: Austria

Age: 31

Place of birth: Vienna

Education

1995 – 1998	Study to be a Pharmacological Commercial Assistant
1998 – 2000	Graduate High School
2000 – 2005	Study of Biology, University of Vienna
	Major: Zoology
	Specific field: Cognition, Behavioural biology
	Minor: Anthropology
	Specific field: Evolutionary Psychology
	Grading: Passed with distinction

2005	Diploma thesis “Tool selection in a nut-cracking task by capuchin monkeys (*Cebus apella*)”
	Supervisors: Dr. Elisabetta Visalberghi, CNR, Rome Ao. Univ.-Prof. Mag. Dr. Ludwig Huber, University of Vienna
2006	Diploma examination **Grading:** passed with distinction
2007 – present	Doctoral Student “Reasoning about weight in great apes and 3-to-4-year old children”
	Supervisors: Dr. Ludwig Huber, Department of Cognitive Biology, University of Vienna Dr. Josep Call, Max-Planck-Institute for Evolutionary Anthropology, Leipzig
2007 – 2009	Visiting Scientist Institute for Psychology (University of Heidelberg) advised by Dr. Sabina Pauen Great Ape Research Institute Hayashibara, Japan, advised by Dr. Satoshi Hirata

<u>Additional Education</u>

2008 - 2009	Mentoring Program for Women Academics and Scientists at the University of Vienna
2009	Workshop Training für Doktorandinnen: Einstieg in die Lehre an der Universität Wien
2009	Medientraining für Wissenschafterinnen gefördert vom BMWF

<u>Professional history</u>

2006	Employment at the Colobus Trust: Primate Rescue, Research and Conservation Centre, Diani Beach, Kenya.

- Supervised and supported ongoing research projects
- Field work (using GPS) and report writing

Project study: A step too far?

2004 Tropical biology field course in Kibale Nationalpark, Uganda
(organized by Tropical Biology Association, Cambridge).

Project study: The significance of allo-mothering in black and white colobus monkeys (*Collobus guereza*).

2003 Work Placement at the Max PIanck Institute for Evolutionary Anthropology, Department of Comparative Psychology, Leipzig.

- Data collection on chimpanzees and gorillas in a reversed contingency task.

2002 Volunteer at the Zoological garden Zagreb, Primate Section.

- Responsible for basic maintenance and care of primates e.g., planning of enrichment tasks.

2001 – 2003 Student assistant, University of Vienna, Department of Cognitive Biology, Marmosets' lab.

- Assisted in data collection, video analysis, hand rearing of infant marmosets, maintenance and care of marmosets.

<u>Fellowships – International Experience</u>

01/2007 – 01/2010 Recipient of a DOC-fFORTE-fellowship of the Austrian Academy of Sciences,
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany;
Institute for Psychology, University of Heidelberg;
Great Apes Research Centre, Tamano, Japan

10/2006 – 12/2006 Grant from the Austrian Ministry of Science (BMWF)
"Kurzfristige wissenschaftliche Arbeiten im Ausland",
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

10/2006 Würdigungspreis des Bundesministeriums für Bildung, Wissenschaft und Kultur (BMWF) für herausragende Leistungen der Diplomabsolventen und -absolventinnen der Fakultät Lebenswissenschaften

2006 Förderpreis (Dr. Maria Schaumayer Stiftung, Austria)

04/2005 – 06/2005 Grant from the Austrian Ministry of Science (BMWF)
"Kurzfristige wissenschaftliche Arbeiten im Ausland",
Istituto di Scienze e Tecnologie della Cognizione, CNR, Rome, Italy

Professional societies

2010 International Primatological Society (IPS)

2010 Gesellschaft für Primatologie e.V. (GfP)

Publications

*Schrauf C, Fuwa K, Call J, Hirata S (in prep) "Do chimpanzees use weight to select hammer tools?"

*Schrauf C, Call J (submitted) "Great apes can use weight as cue to find hidden food"

*Schrauf C, Call J, Pauen S (resubmitted) "Do 3-4 year old children differ in their expectancies after experiencing physical possible and impossible feedback on a balance scale?"

*Schrauf C, Call J (2009) "Great apes' performance in discriminating weight and achromatic color" Anim Cog *12,* 567-574

Schrauf C, Huber L, Visalberghi E (2008) "Do capuchin monkeys (*Cebus apella*) use weight to select hammer tools?" Anim Cog 11, 413- 422

Machatschke IH, Bauer BE, Schrauf C, Dittami J, Wallner B (2008) "Conflict-involvement of Male Guinea Pigs (*Cavia aperea f. porcellus*) as a Criterion for Partner Preference" Behav Ecol Sociobiol 62, 1341 – 1350

Voelkl B, Schrauf C, Huber L (2006) "Social contact influences the response towards novel food in infant marmosets" Anim Behav 72, 365 -372

*asterisk indicates publications in the frame of my Dissertation

Published abstracts

Schrauf C, Oroni Iseren E, Sebulime P (2005) The significance of allo-mothering in black and white colobus (*Collobus guereza*) in Kibale Forest. Abstracts from TBA Uganda project reports on mammals 1998-2005.

Schrauf C, Voelkl B, Huber L (2004) The response of infant marmosets towards novel food. Folia Primatologia, Vol. 75 Suppl. (XX Congress of the IPS), p. 410-411.

Conference Presentations (Poster)

Schrauf Cornelia, Voelkl Bernhard, Call Josep (2008) "Great apes` performance in discriminating weight and brightness" from the 22th Congress of the International Primatological Society IPS 2008 Edinburgh.

Schrauf Cornelia, Call Josep, Krist Horst, Pauen Sabina (2007) "Do 3-4 year old children differ in their expectancies after experiencing physical possible and impossible feedback on a balance scale?" Poster presented at the Fachtagung für Entwicklungspsychologie, Heidelberg.

Cooper Felicty, Schrauf Cornelia (2006) „A step too far?" Poster presented at the Gesellschaft für Primatologie, 10. Tagung, Leipzig.

Schrauf Cornelia, Voelkl Bernhard, Huber Ludwig (2004) "The response towards novel food in infant common marmosets“ from the 20th Congress of the International Primatological Society IPS 2004. Folia Primatologica. 75, S1, 04, 712, pp. 410 – 411.

Conference Presentations (Talk)

Schrauf Cornelia (2006) “Tool selection in a nut-cracking task in capuchin monkeys (*Cebus apella*)” Talk given at the Gesellschaft für Primatologie, 10. Tagung, Leipzig.

Invited Presentation (Talk)

Schrauf Cornelia (2009) “Do 3-4 year old children differ in their expectancies after experiencing physical possible and impossible feedback on a balance scale?” University of Bristol, Department of Experimental Psychology, Seminar.

Schrauf Cornelia (2008) “Great apes` performance in discriminating weight and brightness” Talk presented at University of Cambridge, Madingley Seminars.

Schrauf Cornelia (2006) “Reasoning about weight in great apes and human children” Talk presented at the Institut für Psychologie, Universität Greifswald, Germany.